

## DISK CONTROLLER

Patent Number: JP2000056932  
Publication date: 2000-02-25  
Inventor(s): TAMAI MASARO; KOKADO TAKESHI; ITO YUKIKO  
Applicant(s): MATSUSHITA ELECTRIC IND CO LTD  
Requested Patent: JP2000056932  
Application Number: JP19980226221 19980810  
Priority Number(s):  
IPC Classification: G06F3/06; G11B19/02  
EC Classification:  
Equivalents:

---

### Abstract

---

**PROBLEM TO BE SOLVED:** To provide a disk controller which satisfies the requested delay time according to the deterioration of performance of a disk device and also executes a request with high efficiency.

**SOLUTION:** A request management part 13 holds the requests as a queue, and a delay time management part 12 holds the end time when the execution should be ended for each request included in the queue. An access time management part 15 holds the access time that is acquired by giving a request to a hard disk 16. A rearrangement control part 11 rearranges the requests every time a request is inputted so as to shorten the seek distance and to end the execution of a request before its end time. A disk IO part 14 measures the execution time of each request and updates the access time. Thus, it's possible to execute a request before its end time according to the deterioration of performance of the disk 16 due to its aging and also to secure the real time nature of animation and voice data.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-56932

(P2000-56932A)

(43) 公開日 平成12年2月25日(2000.2.25)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テマコード(参考)
G 0 6 F 3/06	3 0 2	G 0 6 F 3/06	3 0 2 E 5 B 0 6 5
G 1 1 B 19/02	5 0 1	G 1 1 B 19/02	5 0 1 C

審査請求 未請求 請求項の数32 O L (全 23 頁)

(21) 出願番号 特願平10-226221

(22) 出願日 平成10年8月10日(1998.8.10)

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 玉井 昌朗

大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

(72) 発明者 古門 健

大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

(74) 代理人 100098291

弁理士 小笠原 史朗

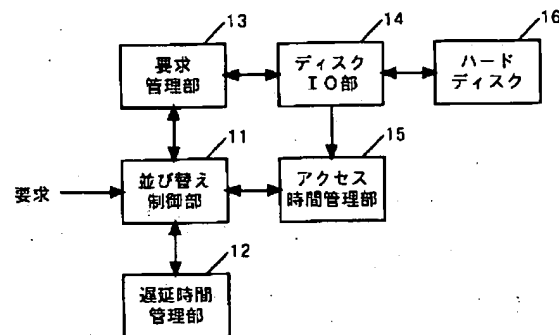
最終頁に続く

(54) 【発明の名称】 ディスク制御装置

(57) 【要約】

【課題】 ディスク装置の性能劣化に対応しつつ、要求された遅延時間を満たし、かつ、効率良く要求を実行するディスク制御装置を提供することである。

【解決手段】 要求管理部13は、要求をキューとして保持する。遅延時間管理部12は、キュー内の要求ごとに要求の実行が完了すべき時刻である完了時刻を保持する。アクセス時間管理部15は、ハードディスク16に要求を発行することで得られたアクセス時間を保持する。並び替え制御部11は、要求が入力されるたびに、シーク距離が短く、かつ完了時刻以前に要求の実行が完了するように要求を並び替える。ディスクI/O部14は、要求の実行時間を測定し、アクセス時間を更新する。これにより、ディスクの経年変化による性能劣化に対応しつつ、完了時刻までに要求実行が可能となり、動画や音声データのリアルタイム性を確保することができる。



## 【特許請求の範囲】

【請求項1】 各要求が個別に有する遅延可能時間以内に当該各要求の実行が完了するように要求を並べ替えてディスク装置に発行するディスク制御装置であって、一つまたは複数の前記要求を保持する要求管理手段と、前記要求管理手段が保持する要求を取り出して前記ディスク装置に発行し、また前記ディスク装置から当該要求に対する応答を受信するディスクI/O手段と、前記ディスク装置における前記各要求の処理に必要な時間をアクセス時間として保持するアクセス時間管理手段と、前記各要求ごとに、要求の処理を完了すべき完了時刻を保持する遅延時間管理手段と、前記完了時刻および前記アクセス時間に基づいて、前記要求管理手段が保持する要求を並べ替える並べ替え制御手段とを備える、ディスク制御装置。

【請求項2】 前記アクセス時間は、前記ディスク装置に要求を発行し、前記ディスク装置から当該要求に対する応答を受信するまでの時間に基づいて定めることを特徴とする、請求項1に記載のディスク制御装置。

【請求項3】 前記アクセス時間は、要求がアクセス可能な前記ディスク装置上の領域を、連続した領域である複数のグループに分割し、前記各要求について要求がアクセスする前記ディスク装置上の位置が属するグループをそれぞれ求め、前記各要求の実行順序に対応した前記グループの順序により定めることを特徴とする、請求項2に記載のディスク制御装置。

【請求項4】 前記アクセス時間は、前記要求の種類ごとに個別に定めることを特徴とする、請求項3に記載のディスク制御装置。

【請求項5】 前記アクセス時間は、前記要求がアクセスするデータサイズごとに個別に定めることを特徴とする、請求項3または4に記載のディスク制御装置。

【請求項6】 前記データサイズは、固定長であることを特徴とする、請求項5に記載のディスク制御装置。

【請求項7】 前記ディスク装置上の領域は、複数の固定長のブロックから構成されており、前記グループは、それぞれ均等な数の前記ブロックを有することを特徴とする、請求項3に記載のディスク制御装置。

【請求項8】 前記ディスク装置は、複数のシリンダで構成されるハードディスクであり、前記グループは、それぞれ均等な数の前記シリンダを有することを特徴とする、請求項3に記載のディスク制御装置。

【請求項9】 前記グループの順序は、2つのグループの順列または組合せであることを特徴とする、請求項3～8のいずれかに記載のディスク制御装置。

【請求項10】 前記要求の種類は、前記ディスク装置からデータを読み出す読み出し要求と、前記ディスク装

置にデータを書き込む書き込み要求であることを特徴とする、請求項4に記載のディスク制御装置。

【請求項11】 前記完了時刻は、前記要求が入力された時刻に当該要求が有する前記遅延可能時間を加えた時間であり、

前記並べ替え制御手段は、前記各要求の実行が各々の前記完了時刻以前に完了するように、各要求を並び替えることを特徴とする、請求項1に記載のディスク制御装置。

【請求項12】 前記遅延時間管理手段は、予め定めた定数を有しており、前記完了時刻は、前記要求が入力された時刻に前記定数を加えた時間であり、

前記並べ替え制御手段は、前記各要求の実行が各々の前記完了時刻以前に完了するように、各要求を並び替えることを特徴とする、請求項1に記載のディスク制御装置。

【請求項13】 前記ディスクI/O手段は、要求が前記ディスク装置に発行してから、前記ディスク装置から当該要求に対する応答を受信するまでの実行時間を測定する実行時間測定手段をさらに備え、

前記アクセス時間管理手段は、前記実行時間を新たなアクセス時間として更新するアクセス時間算出手段をさらに備えることを特徴とする、請求項3～6のいずれかに記載のディスク制御装置。

【請求項14】 前記アクセス時間算出手段は、前記実行時間を前記グループの順序、前記要求の種類および前記データサイズごとに分類し、当該分類に対応するアクセス時間の更新を行うことを特徴とする、請求項13に記載のディスク制御装置。

【請求項15】 前記アクセス時間管理手段は、一定数の要求の実行が完了するごとに、前記分類の各々についてアクセス時間の更新を行うことを特徴とする、請求項13または14に記載のディスク制御装置。

【請求項16】 前記アクセス時間管理手段は、前記分類の各々について一定数の要求の実行が完了するごとに、前記実行が完了した分類のアクセス時間の更新を行うことを特徴とする、請求項13または14に記載のディスク制御装置。

【請求項17】 前記アクセス時間は、分類された前記実行時間の平均値により求めることを特徴とする、請求項13～16のいずれかに記載のディスク制御装置。

【請求項18】 前記ディスクI/O手段は、前記ディスク装置に発行する要求を生成するコマンド生成手段をさらに備え、

前記アクセス時間算出手段は、前記アクセス時間の初期値を求めることを特徴とする、請求項13～17のいずれかに記載のディスク制御装置。

【請求項19】 前記アクセス時間を書き込みコマンドにより記憶し、読み出しコマンドにより出力する記憶装置と、

書き込みコマンドおよび読み出しコマンドを記憶装置に発行するコマンド発行手段とをさらに備え、

前記アクセス時間管理手段は、前記書き込みコマンドを生成して前記アクセス時間を前記記憶装置に格納し、また前記読み出しコマンドを生成して前記アクセス時間を前記記憶装置から取得することを特徴とする、請求項1～18のいずれかに記載のディスク制御装置。

【請求項20】 前記記憶装置は、前記ディスク装置に構成され、前記コマンド発行手段は、前記ディスクI/O手段に構成されることを特徴とする、請求項19に記載のディスク制御装置。

【請求項21】 前記要求管理手段は、各要求ごとに、前記ディスク装置をシークする方向が変化することを示すフラグと、要求が入力された時刻にキュー内で他の要求が完了するのを待つ時間と、前記要求のアクセス時間とを加えた第1の時刻を保持しており、

前記並べ替え制御手段に入力された要求を第1の要求とし、前記要求管理手段内の前記フラグが設定された最後尾の要求、または前記フラグが設定されていなければ前記ディスクI/O手段が前記要求管理手段から取り出した最後尾の要求を第2の要求とした場合、

前記並べ替え制御手段は、

前記第1の要求を、前記第1の要求がアクセスする論理アドレスが、前記要求管理手段内の最後尾の要求から前記第2の要求までがアクセスする各論理アドレスの順序を乱さない位置に仮挿入し、

前記仮挿入により論理アドレスの順序が変更された場合には、前記第1の要求の前記フラグを設定し、

前記仮挿入後に、前記要求管理手段内で前記第1の要求以降の各要求について前記第1の時刻を再び算出し、

前記再び算出した第1の時刻が前記完了時刻を越えない場合は、前記仮挿入した位置を前記第1の要求の挿入位置とし、

前記再び算出した第1の時刻が前記完了時刻を越える場合は、前記要求管理手段の最後尾を前記第1の要求の挿入位置とすることを特徴とする、請求項1に記載のディスク制御装置。

【請求項22】 前記第1の要求と同じデータをアクセスする要求を第3の要求とした場合、

前記並べ替え制御手段は、前記第1の要求と前記第3の要求とは読み出し要求であり、かつ前記第1の要求と前記第3の要求との間に前記第1または第3の要求がアクセスするデータに書き込む他の要求がない場合には、前記第1の要求のアクセス時間を0とし、

前記ディスクI/O手段は、前記第3の要求の実行結果を前記第1の要求の実行結果とすることを特徴とする、請求項21に記載のディスク制御装置。

【請求項23】 前記第1の要求と同じ位置にあるデータに書き込む要求を第3の要求とした場合、

前記並べ替え制御手段は、前記第1の要求は書き込み要

求であり、かつ前記第1の要求と前記第3の要求との間に前記第1または第3の要求がアクセスするデータを読み出す他の要求がない場合には、前記第1の要求を前記第3の要求の直後に挿入すると共に、前記第3の要求のアクセス時間を0とし、

前記ディスクI/O手段は、前記第3の要求を発行せず、前記第1の要求の実行結果を前記第3の要求の実行結果とすることを特徴とする、請求項21に記載のディスク制御装置。

【請求項24】 前記第1の要求が読み出すデータと連続した位置にあるデータを読み出す要求を第3の要求とした場合、

前記並べ替え制御手段は、

前記第1の要求と前記第3の要求との間に前記第1の要求がアクセスするデータに書き込む他の要求がない場合には、前記第1の要求を前記連続した位置の順序に従って前記第3の要求の直前または直後に挿入し、

前記第1の要求または前記第3の要求のアクセス時間を、前記第1の要求と前記第3の要求とが読み出すデータサイズに応じたアクセス時間とし、

前記ディスクI/O手段は、前記第1および第3の要求をまとめ、前記連続した位置にあるデータを読み出す一つの要求として前記ディスク装置に発行することを特徴とする、請求項21に記載のディスク制御装置。

【請求項25】 前記第1の要求が書き込むデータと連続した位置にあるデータを書き込む要求を第3の要求とした場合、

前記並べ替え制御手段は、

前記第1の要求と前記第3の要求との間に前記第1の要求がアクセスするデータに読み出す他の要求がない場合には、前記第1の要求を前記連続した位置の順序に従って前記第3の要求の直前または直後に挿入し、

前記第1の要求または前記第3の要求のアクセス時間を、前記第1の要求と前記第3の要求とが読み出すデータサイズに応じたアクセス時間とし、

前記ディスクI/O手段は、前記第1および第3の要求をまとめ、前記連続した位置にあるデータを読み出す一つの要求として前記ディスク装置に発行することを特徴とする、請求項21に記載のディスク制御装置。

【請求項26】 各要求が個別に有する遅延可能時間以内に当該各要求の実行が完了するように要求を並べ替えてディスク装置に発行するディスク制御装置であって、一つまたは複数の前記要求を保持する要求管理手段と、前記要求管理手段が保持する要求を取り出して前記ディスク装置に発行し、また前記ディスク装置から当該要求に対する応答を受信するディスクI/O手段と、

前記各要求について、前記遅延可能時間より、要求を受け付けた際にすでに前記要求管理手段に保持されていた要求の数に加え、さらに先に実行してもよい要求数を示す挿入可能要求数を求める遅延時間管理手段と、

前記挿入可能要求数に基づいて、前記要求管理手段が保持する要求を並べ替える並べ替え制御手段とを備える、ディスク制御装置。

【請求項27】 前記遅延時間管理手段は、要求を前記ディスク装置に発行してから前記ディスク装置より当該要求に対する応答を受信するまでの時間である基準実行時間を保持する基準実行時間保持手段と、前記基準実行時間と前記各要求内に指定された前記遅延可能時間とから、前記挿入可能要求数を求める挿入可能要求数算出手段とをさらに備えることを特徴とする、請求項26に記載のディスク制御装置。

【請求項28】 前記ディスクIO手段は、前記ディスク装置に発行する要求を生成するコマンド生成手段と、要求が前記ディスク装置に発行してから前記ディスク装置より当該要求に対する応答を受信するまでの実行時間を測定する実行時間測定手段と、前記実行時間測定手段が測定した前記実行時間から前記基準実行時間の初期値を求める基準実行時間算出手段とをさらに備えることを特徴とする、請求項27に記載のディスク制御装置。

【請求項29】 前記ディスクIO手段は、要求が前記ディスク装置に発行してから前記ディスク装置より当該要求に対する応答を受信するまでの実行時間を測定する実行時間測定手段と、前記実行時間測定手段が測定した前記実行時間から新たな前記基準実行時間を求め、前記基準実行時間保持手段が保持する前記基準実行時間の更新を行う基準実行時間算出手段とをさらに備えることを特徴とする、請求項27に記載のディスク制御装置。

【請求項30】 前記基準実行時間は、前記測定した実行時間の平均値により求めることを特徴とする、請求項27～29のいずれかに記載のディスク制御装置。

【請求項31】 前記要求管理手段は、各要求ごとに、前記ディスク装置をシークする方向が変化することまたは前記挿入可能要求数が0であることを示すフラグと、前記挿入可能要求数とを保持しており、前記並べ替え制御手段に入力された要求を第1の要求とし、前記要求管理手段内の前記フラグが設定された最後尾の要求、または前記フラグが設定されていなければ前記ディスクIO手段が前記要求管理手段から取り出した最後尾の要求を第2の要求とした場合、前記並べ替え制御手段は、前記第1の要求を、前記第1の要求がアクセスする論理アドレスが、前記要求管理手段内の最後尾の要求から前記第2の要求までがアクセスする各論理アドレスの順序を乱さない位置に挿入し、前記挿入により論理アドレスの順序が変更された場合には、前記第1の要求の前記フラグを設定し、前記第1の要求の挿入位置以前にある要求数を、前記第

1の要求の前記挿入可能要求数から減算し、前記第1の要求の挿入位置以降にあるすべての要求の前記挿入可能要求数から1を減算し、前記減算により前記挿入可能要求数が0となった要求に対して、前記フラグを設定することを特徴とする、請求項26に記載のディスク制御装置。

【請求項32】 前記ディスク装置がキャッシュ機能を有している場合、前記ディスクIO手段は、前記キャッシュ機能を使用せずに、前記ディスク装置への前記要求の発行、および当該要求に対する応答の受信を行うことを特徴とする、請求項1～31のいずれかに記載のディスク制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、一定の速度で発生する要求をリアルタイムに処理することを必要とするメディアサーバに使用されるディスク制御装置に関し、より特定的には、ランダムに発生する要求に対してシーク時間を減少させて要求の処理能力を向上させたディスク制御装置に関する。

【0002】

【従来の技術】近年、音声データや映像データを含むマルチメディアデータを記憶するディスク装置を備え、通信回線を介して接続された複数のユーザに対してデータを供給したり、データをディスク装置に書き込むメディアサーバが実用化されている。ユーザは、メディアサーバから供給されるデータにより、見たい時に希望する映像を見ることができ、聞きたい時に希望する音声聞くことができる。同様に、ユーザは、メディアサーバにデータを供給することで、映像や音声データをメディアサーバに記録することができる。

【0003】このメディアサーバは、複数のユーザに対してデータをリアルタイムに出力する必要がある。なぜならば、データの出力のタイミングが、予め定められている許容遅延時間を越えてしまうと、音声途切れたり映像が乱れたり（欠落）するという問題が生じるからである。同様に、メディアサーバは、複数のユーザからのデータをリアルタイムに記録する必要がある。なぜならば、端末から音声や映像データをリアルタイムでデジタルデータに変換し、メディアサーバに格納する場合、メディアサーバに格納する処理の完了が、予め定められた許容遅延時間を越えてしまうと、端末側もしくはメディアサーバ側のメモリが溢れてしまうという問題が生じるからである。

【0004】この問題に対応すべく、従来から、メディアサーバにおいてデータをリアルタイムにディスクから読み出すためのディスク制御装置として、特開平8-55055号公報（以下、従来の文献という）に開示されているものがある。以下、図17を用いて、この従来の文献に記載されているディスク制御装置を簡単に説明す

る。

【0005】図17は、この従来の文献に記載されているディスク制御装置の構成を示すブロック図である。図17において、従来のディスク制御装置は、並べ換え制御部301と、許容遅延データ管理部302と、要求管理部303と、ハードディスク306とを備える。

【0006】ハードディスク306は、複数種類の映像や音声を表すマルチメディアデータを一定のサイズごとに分割して記憶する。許容遅延データ管理部302は、各読み出し要求ごとに、当該要求が指示するデータの読み出しが完了するまでに許容される遅延時間を記憶する。要求管理部303は、各読み出し要求を保持する。並べ換え制御部301は各読み出し要求について、読み出しが完了するまでの時間を算出して判断することで、許容遅延データ管理部302が保持する許容遅延時間以内に読み出しが完了し、かつ、ハードディスク306でのシーク距離が短くなるように読み出し要求を並べ替える。

【0007】読み出しが完了するまでの時間は、キュー内での待ち時間と、要求の実行時間との和で算出される。キュー内での待ち時間は、キューにおいて、待ち時間を算出する要求よりも前に格納されている要求の実行時間の和として算出される。要求の実行時間は、その要求がアクセスするデータのシリンダ位置から算出したシークに必要な時間と、ハードディスクの回転数から算出した平均回転待ち時間と、1シリンダあたりのブロック数とハードディスクの回転数とから算出したデータ転送時間との和として求められる。このように、読み出し要求の実行時間は、製品マニュアルなどに記載された値から算出される。このように、従来のディスク制御装置は、要求の完了時間に基づいて各読み出し要求を並び替えることで、処理のリアルタイム性を向上させている。

【0008】

【発明が解決しようとする課題】しかしながら、上記従来の文献に記載されているディスク制御装置では、シーク時間と平均回転待ち時間とデータ転送時間とを算出し、これらの和を読み出し要求の実行時間として求めるため、製品マニュアルなどに記載された値が同じ場合には算出する実行時間は等しくなる。このため、ディスク面上にある不良セクタ数やディスク内の制御ICの動作の違いに基づく実行時間の違いまでは、算出された実行時間に反映されなかった。また、長期間に渡って同じディスクを使用した場合に発生するディスクの性能劣化の影響についても、実行時間の算出に反映されなかった。

【0009】これにより、従来のディスク制御装置においては、算出した実行時間から許容遅延時間以内に要求を実行できるよう並べ替えを行ってはいけるが、実際に要求を実行した場合に、要求の実行完了が許容遅延時間を越えてしまうということが発生していた。

【0010】それ故、本発明の目的は、連続的に入力さ

れる複数の要求に対して、要求の順序を並べ替えることによりシーク時間を短縮し、かつ、許容遅延時間以内に要求の実行を完了するディスク制御装置を提供することである。

【0011】

【課題を解決するための手段および発明の効果】第1の発明は、各要求が個別に有する遅延可能時間以内に当該各要求の実行が完了するように要求を並べ替えてディスク装置に発行するディスク制御装置であって、一つまたは複数の要求を保持する要求管理手段と、要求管理手段が保持する要求を取り出してディスク装置に発行し、またディスク装置から当該要求に対する応答を受信するディスクIO手段と、ディスク装置における各要求の処理に必要な時間をアクセス時間として保持するアクセス時間管理手段と、各要求ごとに、要求の処理を完了すべき完了時刻を保持する遅延時間管理手段と、完了時刻およびアクセス時間に基づいて、要求管理手段が保持する要求を並べ替える並べ替え制御手段とを備える。

【0012】上記のように、第1の発明によれば、要求が入力されるたびに、アクセス時間管理手段が保持するアクセス時間により要求の実行時間を予測し、各要求内に指定された遅延可能時間以内に要求の実行が完了するように要求を並べ替える。要求の実行時間の予測は保持されたアクセス時間を参照するだけであるから、並べ替えの際に予測値を得るために特別な演算を必要としない。これにより、要求が入力されるたびに要求の並べ替えを行なうので、アクセス効率がよくなり、かつ、遅延可能時間以内に要求を実行することが可能となる。特に、動画や音声データなどのリアルタイム性を必要とするデータにアクセスする場合に、アクセス効率を向上させ、かつ、リアルタイム性を確保することが可能となる。

【0013】第2の発明は、第1の発明において、アクセス時間は、ディスク装置に要求を発行し、ディスク装置から当該要求に対する応答を受信するまでの時間に基づいて定めることを特徴とする。

【0014】上記のように、第2の発明によれば、第1の発明において、ディスク装置に要求を発行し、当該要求に対する応答を受信するまでの時間を測定した値からアクセス時間を求めるため、接続したディスク装置の性能に応じた値を求めることができる。

【0015】第3の発明は、第2の発明において、アクセス時間は、要求がアクセス可能なディスク装置上の領域を、連続した領域である複数のグループに分割し、各要求について要求がアクセスするディスク装置上の位置が属するグループをそれぞれ求め、各要求の実行順序に対応したグループの順序により定めることを特徴とする。

【0016】上記のように、第3の発明によれば、第2の発明において、アクセス時間は、各要求がアクセスす

る位置が属するグループの順序により定めることができる。

【0017】第4の発明は、第3の発明において、アクセス時間は、要求の種類ごとに個別に定めることを特徴とする。

【0018】上記のように、第4の発明によれば、第3の発明において、アクセス時間は、各要求の種類によっても定めることができ、要求の種類ごとのより正確な予測が可能となる。

【0019】第5の発明は、第3および第4の発明において、アクセス時間は、要求がアクセスするデータサイズごとに個別に定めることを特徴とする。

【0020】上記のように、第5の発明によれば、第3および第4の発明において、要求がアクセスするデータサイズによっても定めることができ、データサイズごとのより正確な予測が可能となる。

【0021】第6の発明は、第5の発明において、データサイズは、固定長であることを特徴とする。

【0022】上記のように、第6の発明によれば、第5の発明において、アクセス時間の保持に必要なメモリ量を減少することができる。

【0023】第7の発明は、第3の発明において、ディスク装置上の領域は、複数の固定長のブロックから構成されており、グループは、それぞれ均等な数のブロックを有することを特徴とする。

【0024】第8の発明は、第3の発明において、ディスク装置は、複数のシリンドラで構成されるハードディスクであり、グループは、それぞれ均等な数のシリンドラを有することを特徴とする。

【0025】上記のように、第7および第8の発明によれば、第3の発明において、各グループごとの実行時間の揺らぎが少なくなり、より正確な予測値を求めることができる。

【0026】第9の発明は、第3～第8の発明において、グループの順序は、2つのグループの順列または組合せであることを特徴とする。

【0027】上記のように、第9の発明によれば、第3～第8の発明において、2つのグループからなる順列または組合せにより実行時間を求めることができ、予測値の精度を変えずに、より少ないメモリ量での予測が可能となる。

【0028】第10の発明は、第4の発明において、要求の種類は、ディスク装置からデータを読み出す読み出し要求と、ディスク装置にデータを書き込む書き込み要求であることを特徴とする。

【0029】上記のように、第10の発明によれば、第4の発明において、並べ替えを行なう要求を読み出し要求と書き込み要求に限定することで、リアルタイム性を確保すべきデータのみについて、アクセス時間を保持して並べ替えを行なうことができる。

【0030】第11の発明は、第1の発明において、完了時刻は、要求が入力された時刻に当該要求が有する遅延可能時間を加えた時間であり、並べ替え制御手段は、各要求の実行が各々の完了時刻以前に完了するように、各要求を並び替えることを特徴とする。

【0031】上記のように、第11の発明によれば、第1の発明において、各要求の完了時刻を常に更新する必要がなく、演算量を減少させることができる。

【0032】第12の発明は、第1の発明において、遅延時間管理手段は、予め定めた定数を有しており、完了時刻は、要求が入力された時刻に定数を加えた時間であり、並べ替え制御手段は、各要求の実行が各々の完了時刻以前に完了するように、各要求を並び替えることを特徴とする。

【0033】上記のように、第12の発明によれば、第1の発明において、各要求に遅延可能時間を付加する必要がなく、また、完了時刻の算出も簡単化される。

【0034】第13の発明は、第3～第6の発明において、ディスクIO手段は、要求がディスク装置に発行してから、ディスク装置から当該要求に対する応答を受信するまでの実行時間を測定する実行時間測定手段をさらに備え、アクセス時間管理手段は、実行時間を新たなアクセス時間として更新するアクセス時間算出手段をさらに備えることを特徴とする。

【0035】上記のように、第13の発明によれば、第3～第6の発明において、要求の実行時間を測定しつつ、測定した実行時間に基づいてアクセス時間を更新するため、経時によるディスクの性能劣化に対応したアクセス時間を求めることが可能となる。

【0036】第14の発明は、第13の発明において、アクセス時間算出手段は、実行時間をグループの順序、要求の種類およびデータサイズごとに分類し、当該分類に対応するアクセス時間の更新を行うことを特徴とする。

【0037】上記のように、第14の発明によれば、第13の発明において、グループの順序、要求の種類およびデータサイズについてアクセス時間を更新するため、経時によるディスクの性能劣化に対応したアクセス時間を求めることが可能となる。

【0038】第15の発明は、第13および第14の発明において、アクセス時間管理手段は、一定数の要求の実行が完了するごとに、分類の各々についてアクセス時間の更新を行うことを特徴とする。

【0039】第16の発明は、第13および第14の発明において、アクセス時間管理手段は、一定数の要求の実行が完了するごとに、分類の各々について実行が完了した分類のアクセス時間の更新を行うことを特徴とする。

【0040】上記のように、第15および第16の発明によれば、第13および第14の発明において、常に一

定数の要求ごとにアクセス時間が更新されるため、使用頻度に応じて経時によるディスクの性能劣化に対応したアクセス時間を求めることができる。

【0041】第17発明は、第13～第16の発明において、アクセス時間は、分類された実行時間の平均値により求めることを特徴とする。

【0042】上記のように、第17の発明によれば、第13～第16の発明において、偶発的に実行時間が長くなった要求が存在した場合に、求めるアクセス時間に対する影響を抑えることが可能となる。

【0043】第18の発明は、第13～第17の発明において、ディスクIO手段は、ディスク装置に発行する要求を生成するコマンド生成手段をさらに備え、アクセス時間算出手段は、アクセス時間の初期値を求めることを特徴とする。

【0044】上記のように、第18の発明によれば、第13～第17の発明において、接続されたディスク制御装置を使用する際に、アクセス時間の初期値を求めることが可能となる。

【0045】第19の発明は、第1～第18の発明において、アクセス時間を書き込みコマンドにより記憶し、読み出しコマンドにより出力する記憶装置と、書き込みコマンドおよび読み出しコマンドを記憶装置に発行するコマンド発行手段とをさらに備え、アクセス時間管理手段は、書き込みコマンドを生成してアクセス時間を記憶装置に格納し、また読み出しコマンドを生成してアクセス時間を記憶装置から取得することを特徴とする。

【0046】第20の発明は、第19の発明において、記憶装置は、ディスク装置に構成され、コマンド発行手段は、ディスクIO手段に構成されることを特徴とする。

【0047】上記のように、第19および第20の発明によれば、第1～第19の発明において、アクセス時間を記憶装置に記憶しておき、使用時に読み出すことでアクセス時間の初期値を求める動作を簡単化することができる。

【0048】第21の発明は、第1の発明において、要求管理手段は、各要求ごとに、ディスク装置をシークする方向が変化することを示すフラグと、要求が入力された時刻にキュー内で他の要求が完了するのを待つ時間と、要求のアクセス時間とを加えた第1の時刻を保持しており、並べ替え制御手段に入力された要求を第1の要求とし、要求管理手段内のフラグが設定された最後尾の要求、またはフラグが設定されていなければディスクIO手段が要求管理手段から取り出した最後尾の要求を第2の要求とした場合、並べ替え制御手段は、第1の要求を、第1の要求がアクセスする論理アドレスが、要求管理手段内の最後尾の要求から第2の要求までがアクセスする各論理アドレスの順序を乱さない位置に仮挿入し、仮挿入により論理アドレスの順序が変更された場合に

は、第1の要求のフラグを設定し、仮挿入後に、要求管理手段内で第1の要求以降の各要求について第1の時刻を再び算出し、再び算出した第1の時刻が完了時刻を越えない場合は、仮挿入した位置を第1の要求の挿入位置とし、再び算出した第1の時刻が完了時刻を越える場合は、要求管理手段の最後尾を第1の要求の挿入位置とすることを特徴とする。

【0049】上記のように、第21の発明によれば、第1の発明において、各要求がもつ遅延可能時間以内に要求の実行が完了するように、かつアクセスする論理アドレスの順に要求を並べ替えることで、アクセスの効率を向上することができる。

【0050】第22の発明は、第21の発明において、第1の要求と同じデータをアクセスする要求を第3の要求とした場合、並べ替え制御手段は、第1の要求と第3の要求とは読み出し要求であり、かつ第1の要求と第3の要求との間に第1または第3の要求がアクセスするデータに書き込む他の要求がない場合には、第1の要求のアクセス時間を0とし、ディスクIO手段は、第3の要求の実行結果を第1の要求の実行結果とすることを特徴とする。

【0051】上記のように、第22の発明によれば、第21の発明において、同じデータをアクセスする読み出し要求を効率良く実行することができる。

【0052】第23の発明は、第21の発明において、第1の要求と同じ位置にあるデータに書き込む要求を第3の要求とした場合、並べ替え制御手段は、第1の要求は書き込み要求であり、かつ第1の要求と第3の要求との間に第1または第3の要求がアクセスするデータを読み出す他の要求がない場合には、第1の要求を第3の要求の直後に挿入すると共に、第3の要求のアクセス時間を0とし、ディスクIO手段は、第3の要求を発行せず、第1の要求の実行結果を第3の要求の実行結果とすることを特徴とする。

【0053】上記のように、第23の発明によれば、第21の発明において、同じデータをアクセスする書き込み要求を効率良く実行することができる。

【0054】第24の発明は、第21の発明において、第1の要求が読み出すデータと連続した位置にあるデータを読み出す要求を第3の要求とした場合、並べ替え制御手段は、第1の要求と第3の要求との間に第1の要求がアクセスするデータに書き込む他の要求がない場合には、第1の要求を連続した位置の順序に従って第3の要求の直前または直後に挿入し、第1の要求または第3の要求のアクセス時間を、第1の要求と第3の要求とが読み出すデータサイズに応じたアクセス時間とし、ディスクIO手段は、第1および第3の要求をまとめ、連続した位置にあるデータを読み出す一つの要求としてディスク装置に発行することを特徴とする。

【0055】上記のように、第24の発明によれば、第



21の発明において、連続したデータをアクセスする複数の読み出し要求を効率良くアクセスすることができる。

【0056】第25の発明は、第21の発明において、第1の要求が書き込むデータと連続した位置にあるデータを書き込む要求を第3の要求とした場合、並べ替え制御手段は、第1の要求と第3の要求との間に第1の要求がアクセスするデータに読み出す他の要求がない場合には、第1の要求を連続した位置の順序に従って第3の要求の直前または直後に挿入し、第1の要求または第3の要求のアクセス時間を、第1の要求と第3の要求とが読み出すデータサイズに応じたアクセス時間とし、ディスクIO手段は、第1および第3の要求をまとめ、連続した位置にあるデータを読み出す一つの要求としてディスク装置に発行することを特徴とする。

【0057】上記のように、第25の発明によれば、第21の発明において、連続したデータをアクセスする複数の書き込み要求を効率良くアクセスすることができる。

【0058】第26の発明は、各要求が個別に有する遅延可能時間以内に当該各要求の実行が完了するように要求を並べ替えてディスク装置に発行するディスク制御装置であって、一つまたは複数の要求を保持する要求管理手段と、要求管理手段が保持する要求を取り出してディスク装置に発行し、またディスク装置から当該要求に対する応答を受信するディスクIO手段と、各要求について、遅延可能時間より、要求を受け付けた際にすでに要求管理手段に保持されていた要求の数に加え、さらに先に実行してもよい要求数を示す挿入可能要求数を求める遅延時間管理手段と、挿入可能要求数に基づいて、要求管理手段が保持する要求を並べ替える並べ替え制御手段とを備える。

【0059】上記のように、第26の発明によれば、要求が入力されるたびに、要求が保持する遅延可能時間から挿入可能要求数を求め、挿入可能要求数よりも多くの要求をその要求の前に挿入しないように、要求を並べ替える。これにより、アクセスの効率が良くなり、また遅延可能時間以内に要求の実行を完了することができる。

【0060】第27の発明は、第26の発明において、遅延時間管理手段は、要求をディスク装置に発行してからディスク装置より当該要求に対する応答を受信するまでの時間である基準実行時間を保持する基準実行時間保持手段と、基準実行時間と各要求内に指定された遅延可能時間とから、挿入可能要求数を求める挿入可能要求数算出手段とをさらに備えることを特徴とする。

【0061】上記のように、第27の発明によれば、第26の発明において、ディスク装置に発行した要求の実行時間を測定した値である基準実行時間を測定し、基準実行時間と各要求が有する遅延可能時間とから挿入可能要求数を求めることにより、より正確な挿入可能要求数

を得ることができる。

【0062】第28の発明は、第27の発明において、ディスクIO手段は、ディスク装置に発行する要求を生成するコマンド生成手段と、要求がディスク装置に発行してからディスク装置より当該要求に対する応答を受信するまでの実行時間を測定する実行時間測定手段と、実行時間測定手段が測定した実行時間から基準実行時間の初期値を求める基準実行時間算出手段とをさらに備えることを特徴とする。

【0063】上記のように、第28の発明によれば、第27の発明において、ディスク制御装置に接続されたディスク装置に発行する要求を生成し、その要求の実行時間を測定することで、接続されたディスク装置に対する基準実行時間の初期値を求めることが可能となる。

【0064】第29の発明は、第27の発明において、ディスクIO手段は、要求がディスク装置に発行してからディスク装置より当該要求に対する応答を受信するまでの実行時間を測定する実行時間測定手段と、実行時間測定手段が測定した実行時間から新たな基準実行時間を求め、基準実行時間保持手段が保持する基準実行時間の更新を行う基準実行時間算出手段とをさらに備えることを特徴とする。

【0065】上記のように、第29の発明によれば、第27の発明において、要求の実行時間を測定し基準実行時間を更新することで、経時変化によるディスク装置の性能劣化に対応した挿入可能要求数を得ることができる。

【0066】第30の発明は、第27～第29の発明において、基準実行時間は、測定した実行時間の平均値により求めることを特徴とする。

【0067】上記のように、第30の発明によれば、第27～第29の発明において、複数の値の平均値を取ることにより、偶発的に要求の実行時間が長くなった要求が存在した場合に基準実行時間に与える影響を抑えることが可能となる。

【0068】第31の発明は、第26の発明において、要求管理手段は、各要求ごとに、ディスク装置をシークする方向が変化する事または挿入可能要求数が0であることを示すフラグと、挿入可能要求数とを保持しており、並べ替え制御手段に入力された要求を第1の要求とし、要求管理手段内のフラグが設定された最後尾の要求、またはフラグが設定されていなければディスクIO手段が要求管理手段から取り出した最後尾の要求を第2の要求とした場合、並べ替え制御手段は、第1の要求を、第1の要求がアクセスする論理アドレスが、要求管理手段内の最後尾の要求から第2の要求までがアクセスする各論理アドレスの順序を乱さない位置に挿入し、挿入により論理アドレスの順序が変更された場合には、第1の要求のフラグを設定し、第1の要求の挿入位置以前にある要求数を、第1の要求の挿入可能要求数から減算

し、第1の要求の挿入位置以降にあるすべての要求の挿入可能要求数から1を減算し、減算により挿入可能要求数が0となった要求に対して、フラグを設定することを特徴とする。

【0069】上記のように、第31の発明によれば、第26の発明において、各要求が有する遅延可能時間以内に要求の実行が完了するように、かつ、各要求がアクセスする論理アドレス順に要求を並べ替えることで、アクセス効率を向上することが可能となる。

【0070】第32の発明は、第1～第31の発明において、ディスク装置がキャッシュ機能を有している場合、ディスクIO手段は、キャッシュ機能を使用せずに、ディスク装置への要求の発行、および当該要求に対する応答の受信を行うことを特徴とする。

【0071】上記のように、第32の発明によれば、第1～第31の発明において、ディスク装置がキャッシュ機能を有している場合でも、キャッシュ機能を使用せずに要求に対する処理を行う。これにより、いかなるディスク装置においても第1～第31の発明の有用な効果を奏することが可能となる。

【0072】

【発明の実施の形態】(第1の実施形態)図1は、本発明の第1の実施形態に係るディスク制御装置の構成を示すブロック図である。図1において、本第1の実施形態に係るディスク制御装置は、並べ替え制御部11と、遅延時間管理部12と、要求管理部13と、ディスクIO部14と、アクセス時間管理部15と、ハードディスク16とを備える。

【0073】並べ替え制御部11は、ユーザ管理部(図示せず)から処理の要求を受け付ける。そして、並べ替え制御部11は、遅延時間管理部12およびアクセス時間管理部15から与えられる情報に基づいて、シーク距離が短くなるように、要求管理部13が管理する処理待ち要求の並び(キュー)の任意の位置に受け付けた要求を挿入する。

【0074】ここで、ユーザ管理部とは、複数のユーザから特定の映像や音声の読み出しまたは書き込み要求を受けて、当該要求に関するデータの処理要求を順次並べ替え制御部11へ発行するものである。このユーザ管理部が発行する処理の要求は、種類と、論理アドレスと、データサイズと、許容遅延時間とを含んでいる。種類とは、要求が読み出し要求であるか書き込み要求であるかを示すものである。論理アドレスとは、処理すべきデータのハードディスク16上の先頭位置を示す論理的なアドレスである。データサイズとは、処理すべきデータのサイズである。許容遅延時間とは、要求の処理が完了するまでに許容される(例えば、ユーザ側で連続した映像または音声として再生するための)最大遅延時間である。

【0075】遅延時間管理部12は、要求管理部13が

管理する要求の完了時刻をそれぞれ算出して保持する。ここで、完了時刻とは、要求の処理が完了しなければならない時刻を表す。要求管理部13は、並べ替え制御部11から連続的に入力される複数の要求をキューとして保持する。ディスクIO部14は、要求管理部13のキューから要求を順に読み出し、対応するコマンドをハードディスク16へ出力する。このコマンドは、所定のサイズのデータの読み出しまたは書き込みを行うコマンドである。アクセス時間管理部15は、ディスクIO部14がコマンドを発行してからハードディスク16より応答を受信するまでの時間をアクセス時間として保持する。ハードディスク16は、複数種類のマルチメディアデータを一定のサイズごとに分割して記憶する。

【0076】図2は、要求管理部13が管理するキューの模式図を示す。図2において、R(5)、R(10)、R(23)…は、読み出し要求または書き込み要求を表す。キューの先頭の要求R(10)は、ハードディスク16における1つ前の要求R(5)が終了した時点で、ディスクIO部14によって取り出される。また、要求R(x)は、並べ替え制御部11から新たに出力された要求を表す(xは、正の整数。以下同じ)。

【0077】まず、図3～図8を用いて、本第1の実施形態に係るディスク制御装置を具体的に説明する。図3は、本第1の実施形態に係るディスク制御装置のさらに詳細な構成を示すブロック図である。図4は、ハードディスク16におけるブロック分割およびグループ分けの一例を示す図である。図5は、完了時刻テーブル122の構成の一例を示す図である。図6は、要求テーブル131の構成の一例を示す図である。図7は、コマンド管理テーブル141aの構成の一例を示す図である。図8は、アクセス時間テーブル152の構成の一例を示す図である。

【0078】まず、以下の説明において、ハードディスク16は、予め図4に示すように任意のバイト単位(以下、ブロックという)で領域分割されており、連続する複数のブロックごとにグループg(1)～g(m)にグループ分けがされているものとする(mは、正の整数)。なお、本第1の実施形態においては、64Kバイト単位にブロック分割されているものとして説明する。

【0079】ここで、ブロックをグループへ分割する方法の一例としては、各グループが均等なブロック数を持つ分割方法がある。この方法では、ハードディスク16内の全ブロック数を最大グループ数mで割った商の数だけ、各グループにブロックを割り当てる。論理アドレスが先頭のブロックから順に上記商の数だけグループg(1)に割り当て、続きのブロックを同様にグループg(2)に、以下順にグループg(m)まで割り当てていく。また、他の分割方法の一例としては、ハードディスク16上のシリンダ数を最大グループ数mで割った商を求め、各グループに上記商の数のシリンダ数が割り当て

られるよう、先頭のシリンダに属しているブロックから順にブロックを割り当てる方法がある。

【0080】図3を参照して、遅延時間管理部12は、完了時刻算出部121と、完了時刻テーブル122とを備える。また、ディスクIO部14は、要求選択部141と、コマンド発行部142と、コマンド生成部143とを備える。アクセス時間管理部15は、グループ変換部151と、アクセス時間テーブル152と、アクセス時間算出部153とを備える。また、要求管理部13は、キューを要求テーブル131の形態で管理している。

【0081】完了時刻算出部121は、並べ替え制御部11から与えられる許容遅延時間 $\text{delay}(x)$ に基づいて、完了時刻 $f(x)$ を算出する。完了時刻テーブル122は、図5に示すように、完了時刻算出部121が算出した完了時刻 $f(x)$ を各要求 $R(x)$ ごとに記憶する。

【0082】要求テーブル131は、上述のように要求を管理するためのテーブルである。要求テーブル131は、図6に示すように、各要求 $R$ ごとに種類 $k$ 、論理アドレス $a$ 、アクセス時間 $t$ 、完了予定時刻 $e$ 、禁止フラグ $\text{flag}$ およびアクセスタイプ $\text{type}$ をそれぞれ記憶する。なお、図6における順序 $P$ は、要求の処理順序、すなわち、要求選択部141が取り出す要求の順序を示している。1がキューの先頭を、最大値 $n$ がキューの末尾を表す。この $n$ は、キュー長でもある。

【0083】種類 $k$ は、要求が読み出し要求であるか書き込み要求であるかを表す。論理アドレス $a$ は、要求 $R$ によって処理すべきデータのハードディスク16上の先頭位置を示す論理的なアドレスである。アクセス時間 $t$ は、要求 $R$ の（または、要求 $R$ までの）実行時間 $\text{exe}$ を予測した値である。要求 $R$ の実行時間 $\text{exe}$ とは、要求 $R$ が読み出し要求の場合、対応するデータをハードディスク16から読み取って出力するのに要する時間であり、要求 $R$ が書き込み要求の場合、対応するデータをハードディスク16に書き込んで応答を受信するまでに要する時間である。要求テーブル131におけるアクセス時間 $t$ は、アクセス時間テーブル152から与えられる。完了予定時刻 $e$ は、要求 $R$ の実行が完了すると予測した時刻である。禁止フラグ $\text{flag}$ は、フラグONによって要求 $R$ の前に新たな要求を挿入できないこと、または、要求 $R$ の前後でシーク方向が逆になることを表す。アクセスタイプ $\text{type}$ は、この要求 $R$ とその直前に処理される要求 $R$ との関係を示すものであり、「独立」、「同一」、「連続」のいずれかが表される。例えば、要求 $R(x)$ の直前の要求が $R(x_{-1})$ である場合、 $\text{type}(x) = \text{「独立」}$ は、要求 $R(x)$ を単独で実行することとし、 $\text{type}(x) = \text{「同一」}$ は、要求 $R(x)$ が要求 $R(x_{-1})$ と同じ論理アドレス $a(x_{-1})$ にアクセスすることを表し、 $\text{type}(x) = \text{「連$

続」は、要求 $R(x)$ が要求 $R(x_{-1})$ の論理アドレス $a(x_{-1})$ に連続した論理アドレスにアクセスすることを表す。

【0084】要求選択部141は、要求テーブル131の先頭から1つまたは連続して（順序 $P$ の並びで）複数の要求を取り出し、ハードディスク16に発行するコマンド $C$ を生成する。コマンド $C$ の生成にあたり、要求選択部141は、コマンド管理テーブル141aを有しており、当該コマンド管理テーブル141aに取り出した要求 $R$ と生成したコマンド $C$ との関係を登録する。コマンド管理テーブル141aは、図7に示すように、各コマンド $C$ ごとに種類 $K$ 、先頭論理アドレス $A$ 、データサイズ $\text{len}$ 、要求 $R$ およびデータ位置 $d$ をそれぞれ記憶する。

【0085】種類 $K$ は、コマンド $C$ が読み出し処理か書き込み処理かを表すもので、上述の種類 $k$ と同じものを格納する。先頭論理アドレス $A$ は、コマンド $C$ によってアクセスするハードディスク16上の先頭の論理アドレスを示す。データサイズ $\text{len}$ は、どれだけのブロックを連続してアクセスするかを示した値である。データ位置 $d$ は、要求 $R$ が先頭から何番目のブロックにアクセスするかを指定した値で、各要求ごとに設定する。図7に示す最初の例では、 $3 \times 64\text{k}$ バイトの連続したブロックを読み出す3つの要求 $R(1) \sim R(3)$ に対して、1つのコマンド $C(1)$ を生成した場合を示している。この場合、データサイズ $\text{len}(1)$ は3、データ位置 $d(1) \sim d(3)$ は順に1～3、先頭論理アドレス $A(1)$ はアクセスするブロックの先頭の論理アドレスとなる。

【0086】コマンド生成部143は、アクセス時間テーブル152に格納するアクセス時間 $t$ の初期値を得るためのコマンド $C$ を生成する。コマンド発行部142は、要求選択部141またはコマンド生成部143が生成したコマンド $C$ を受けて、ハードディスク16へ発行する。また、コマンド発行部142は、発行したコマンド $C$ の実行結果をハードディスク16から受信し、要求選択部141に実行結果を通知する。さらに、コマンド発行部142は、コマンド発行から実行結果の受信までの実行時間 $\text{exe}$ を測定し、アクセス時間算出部153へ出力する。

【0087】アクセス時間算出部153は、コマンド発行部142から受けた実行時間 $\text{exe}$ に基づいて、アクセス時間テーブル152にアクセス時間 $t$ を格納する。グループ変換部151は、ハードディスク16内におけるグループ分けの情報を予め有している。グループ変換部151は、並べ替え制御部11から論理アドレス $a$ を入力して、当該論理アドレス $a$ が属するグループ $g$ を求める。

【0088】アクセス時間テーブル152は、図8(a)に示すように、要求 $R$ の種類 $k$ と要求 $R$ がアクセ

スするデータサイズ  $len$  とをパラメータとして、上述したハードディスク16内のグループ(図4を参照)相互間のアクセス時間  $t$  を順列でそれぞれ記憶する。図8(a)は、ハードディスク16上の全ブロックを  $m$  個のグループに分割した場合における64kバイトの読み出し要求のアクセス時間テーブルを示している。図8(a)において、例えば  $t\{1, 3\}$  は、グループ  $g(1)$  からグループ  $g(3)$  へアクセスするのに必要な時間を表している。従って、この場合においては、アクセス時間  $t = t\{1, 3\}$  となる。

【0089】なお、このアクセス時間管理テーブル152は、アクセスするデータサイズに対応して種々のテーブルを任意に設定することができる(例えば、128kバイトや256kバイト等)。また、読み出し要求に対応するテーブルと書き込み要求に対応するテーブルとは、別個に作成される。また、このアクセス時間管理テーブル152は、グループ  $g$  の順列を考慮しない場合には、図8(b)のように組合せに対応したテーブルとすることもできる。この場合、格納するデータ量を削減することができる。また、アクセス時間を定めるグループ  $g$  は、3以上の値をとってもよい。3つ以上のグループ  $g$  の順列から定める場合は、3つ以上の要求がアクセスする論理アドレスから求まるグループ  $g$  でアクセス時間を決定する。

【0090】次に、図9～図11を参照して、本第1の実施形態に係るディスク制御装置が行う処理を具体的に説明する。図9は、アクセス時間テーブル152の初期設定手順を示すフローチャートである。図10は、新たな要求  $R(x)$  の挿入処理手順を示すフローチャートである。図11は、図10におけるステップS214の算出処理手順を示すフローチャートである。

【0091】まず、前提として、運用時にアクセス時間テーブル152に格納するアクセス時間  $t$  の初期値を求める方法を、図9を参照して説明する。ここでは、64kバイトの読み出し要求に対応するテーブルを求める場合について説明する。コマンド生成部143は、ハードディスク16内のグループ  $g(1) \sim g(m)$  のそれぞれに関し、グループに属する少なくとも1つ以上のブロックを選択する(ステップS101)。例えば、各グループごとにグループ内の先頭、中央、最後の3つのブロックを選択する。次に、コマンド生成部143は、ステップS101で選択した全てのブロックの中から2つのブロックを選択した場合に生成可能な順列をすべて作成する(ステップS102)。例えば、グループ  $g(1)$  の先頭ブロックとグループ  $g(1)$  の中央ブロック、グループ  $g(1)$  の先頭ブロックとグループ  $g(1)$  の最後ブロック、グループ  $g(1)$  の先頭ブロックとグループ  $g(2)$  の先頭ブロック、…という具合である。次に、コマンド生成部143は、ステップS102で生成したすべての順列に関して、ブロックを順列通りにアク

セスするコマンド  $C$  を生成する(ステップS103)。この各コマンドに指定する先頭論理アドレスは、各ブロックの先頭を指す論理アドレスであり、コマンドの種類は読み出し要求、データサイズは64kバイトである。

【0092】コマンド発行部142は、コマンド生成部143が生成したコマンド  $C$  を受けとり、ハードディスク16へ発行し、その応答を受信するまでの実行時間  $exe$  を測定する(ステップS104)。アクセス時間算出部153は、コマンド発行部142からコマンド  $C$  に従ってアクセスした論理アドレスと実行時間  $exe$  とを取得する。そして、アクセス時間算出部153は、グループ変換部151を参照してコマンド  $C$  がどのグループ間の処理かを求め、当該グループ間のアクセス時間  $t$  としてアクセス時間テーブル152の対応する部分に実行時間  $exe$  を格納する(ステップS105)。以上のステップS103～S105の処理を、ステップS102で生成したすべての順列に対して行う(ステップS106)。

【0093】なお、同一のグループ間に関して実行時間  $exe$  を複数求めた場合(例えば、グループ  $g(1)$  の先頭ブロックからグループ  $g(2)$  の先頭ブロックへの実行時間  $exe$  と、グループ  $g(1)$  の最後ブロックとグループ  $g(2)$  の最後ブロックへの実行時間  $exe$  との双方を求めた場合)には、アクセス時間算出部153は、同一のグループ間に関して実行時間  $exe$  の平均値を算出し、当該平均値をアクセス時間テーブル152に格納するようにすればよい。

【0094】次に、並べ替え制御部11が新しい要求  $R(x)$  を受けとった後の動作を、図10を参照して説明する。並べ替え制御部11は、ユーザ管理部から新たな要求  $R(x)$  を受け取ると、要求  $R(x)$  の許容遅延時間  $delay(x)$  を完了時刻算出部121へ送信する。完了時刻算出部121は、現在の時刻に許容遅延時間  $delay(x)$  を加えることで完了時刻  $f(x)$  を算出する(ステップS201)。完了時刻算出部121は、算出した完了時刻  $f(x)$  を完了時刻テーブル122に格納するとともに、並べ替え制御部11へ返送する。

【0095】また、並べ替え制御部11は、ユーザ管理部から要求  $R(x)$  を受け取ると、要求テーブル131をキューの最後尾から順にスキャンし、要求  $R(x)$  と種類  $k$  および論理アドレス  $a$  が等しい要求  $R$  が存在するか否かを検索する(ステップS202)。この検索処理および後述する比較処理の対象は、要求テーブル131に存在する要求のみならず、直前に要求選択部141によって取り出された要求をも含まれる。ステップS202において等しい要求  $R(h)$  が見つかった場合、並べ替え制御部11は、要求  $R(x)$  を要求  $R(h)$  の直後の位置に仮挿入し、要求  $R(h)$  のアクセスタイプ  $pe(h)$  を「同一」に、要求  $R(x)$  の完了予定時刻

$e(x)$ を「 $e(h)$ と同じ値」に設定する(ステップS203)。一方、ステップS202において等しい要求Rが見つからなかった場合、並べ替え制御部11は、次に、要求テーブル131をキューの最後尾から順にスキャンし、要求R(x)と種類kが等しく、かつ、論理アドレスaが論理アドレスa(x)の1ブロック分だけ前後に設定された要求Rが存在するか否かを検索する(ステップS204)。

【0096】上記ステップS204において該当する要求R(i)が見つかった場合、並べ替え制御部11は、論理アドレスa(x)と論理アドレスa(i)とを比較する(ステップS205)。このステップS205において $a(x) > a(i)$ である場合、並べ替え制御部11は、要求R(x)を要求R(i)の直後の位置に仮挿入し、要求R(i)のアクセスタイプtype(i)を「連続」に設定する(ステップS206)。一方、ステップS205において $a(x) < a(i)$ である場合、並べ替え制御部11は、要求R(x)を要求R(i)の直前の位置に仮挿入し、要求R(x)のアクセスタイプtype(x)を「連続」に設定する(ステップS207)。

【0097】上記ステップS204において該当する要求Rが見つからなかった場合、並べ替え制御部11は、要求テーブル131をキューの最後尾から禁止フラグflagがONである要求Rが存在するか否かを検索する(ステップS208)。ステップS208において該当する要求R(j)が見つかった場合、並べ替え制御部11は、その要求R(j)から再後尾までのすべての要求Rの論理アドレスaと、要求R(x)の論理アドレスa(x)との位置を比較する(ステップS209、S211)。ステップS208において該当する要求Rが見つからなかった場合、並べ替え制御部11は、すべての要求Rの論理アドレスaと、要求R(x)の論理アドレスa(x)との位置を比較する(ステップS210、S211)。

【0098】上記ステップS211において論理アドレスの位置を比較した結果、論理アドレスa(x)が比較した論理アドレスの間に位置する場合は、論理アドレスaの順序を乱さない(昇順または降順となる)位置に要求R(x)を仮挿入し、要求R(x)のアクセスタイプtype(x)を「独立」に設定する(ステップS212)。ここで、ステップS211の比較において、論理アドレスa(x)にアクセスする要求R(q)が見つかった場合、要求R(x)と要求R(q)とは、同じ論理アドレスをアクセスする種類kが異なる要求となる。この場合、実行順序を保存するため、要求R(q)の直後に要求R(x)を仮挿入し、要求R(q)のアクセスタイプtype(q)を「連続」に設定する。一方、ステップS211において論理アドレスの位置を比較した結果、論理アドレスa(x)が比較した論理アドレスの間

に位置しない場合は、キューの最後尾に要求R(x)を仮挿入し、要求R(x)のアクセスタイプtype(x)を「独立」に設定する(ステップS213)。この時、論理アドレスa(x)が最後尾の要求Rの論理アドレスa、または挿入位置直前の要求Rの論理アドレスaの順序と異なる場合には、最後尾または挿入位置直前の要求Rの禁止フラグflagをONに設定する。

【0099】次に、並び替え制御部11は、要求R(x)の完了予定時刻 $e(x)$ を算出する(ステップS214)。この完了予定時刻 $e(x)$ を算出するステップS214で行う処理は、図11にサブルーチンとして示す。なお、以下の説明において、要求R(x)のn個前の要求を要求R( $x_{-n}$ )と、要求R(x)のn個後の要求を要求R( $x_{+n}$ )と表現する。

【0100】図11を参照して、並び替え制御部11は、上記ステップS201～S213までの処理に従って要求R(x)の仮挿入を行った後、参照する要求がいくつ前の要求R( $x_{-n}$ )かを示すカウンタnに1を、どれだけ連続した処理が行えるかを示すカウンタdsに1を代入して初期化する(ステップS301)。次に、並び替え制御部11は、要求R(x)のn個前の要求、すなわち、最初は直前の要求R( $x_{-1}$ )のアクセスタイプtype( $x_{-1}$ )を判断する(ステップS302)。ステップS302の判断でアクセスタイプtype( $x_{-1}$ )が「独立」である場合、データサイズを1(=ds)、先頭の論理アドレスをa( $x_{-1}$ )とする(ステップS303)。ステップS302の判断でアクセスタイプtype( $x_{-1}$ )が「同一」である場合、カウンタnに1を加えて判断する対象を変えてステップS302に戻る(ステップS304)。一方、ステップS302の判断でアクセスタイプtype( $x_{-1}$ )が「連続」である場合、カウンタnに1を加えて判断する対象を変え、と共に、カウンタdsに1を加えてステップS302に戻る(ステップS304、S305)。そして、並び替え制御部11は、上記ステップS302～S305を実行して求めたデータサイズ(=ds)と先頭の論理アドレスa( $x_{-n}$ )を、要求R(x)の種類k(x)および論理アドレスa(x)と共に、グループ変換部151へ送出する。

【0101】なお、上記データサイズ(=ds)の上限は、アクセス時間テーブル152が保持しているテーブルに制限される。例えば、アクセス時間テーブル152が4×64kバイトのテーブルまでしか保持していない場合には、データサイズは4が上限となる。従って、より大きなデータサイズに対応したアクセス時間テーブル152を保持することで、この数値の制限をより大きな値に設定することが可能である。

【0102】グループ変換部151は、並び替え制御部11から入力される論理アドレスa( $x_{-n}$ )およびa(x)に基づいて対応するグループg( $x_{-n}$ )およびg

(x)を算出する(ステップS306)。そして、アクセス時間テーブル152は、種類k(x)とデータサイズとが該当するテーブルより、グループg(x<sub>-n</sub>)からグループg(x)までのアクセス時間t{g(x<sub>-n</sub>), g(x)}を抽出する(ステップS307)。

【0103】上記ステップS307で抽出がされると、

$$\begin{aligned} e(x) &= e(x_{-n}) + t\{g(x_{-n}), g(x)\} \\ &= e(x_{-n}) + t(x) \end{aligned} \quad \dots (1)$$

【0104】次に、並び替え制御部11は、要求R(x)以降の要求R(x<sub>+n</sub>)の完了予定時刻e(x<sub>+n</sub>)を下記式(2)に基づいて更新する(ステップS30

$$\begin{aligned} e(x_{+n}) &= e(x_{+n})' + t(x) + t\{g(x), g(x_{+n})\} \\ &\quad - t\{g(x_{-n}), g(x_{+n})\} \end{aligned} \quad \dots (2)$$

【0105】再び図10を参照して、上述のように完了予定時刻e(x)の算出および完了予定時刻e(x<sub>+n</sub>)の更新が終了すると、並び替え制御部11は、新たに算出したすべての完了予定時刻e(x)およびe(x<sub>+n</sub>)が、完了時刻テーブル122に格納している完了時刻f(x)およびf(x<sub>+n</sub>)以内であるか否かをそれぞれ確認する(ステップS215)。ステップS215の確認においてすべて完了時刻f以内である場合は、要求R(x)の仮挿入位置を本挿入位置として並び替え処理を終了する。一方、ステップS215の確認においてすべて完了時刻f以内でない場合は、仮挿入位置は不適当と判断し、今まで行った各要求Rの完了予定時刻eの変更を仮挿入処理前の値に戻し、要求R(x)をキューの再後尾に挿入し、要求R(x)のアクセスタイプtype(x)を「独立」に設定して、並び替え処理を終了する(ステップS216)。この時、要求R(x)の論理アドレスa(x)がもともと再後尾にあった要求の論理アドレスの順序と異なる場合、当該もともと再後尾にあった要求の禁止フラグflagをONに設定する。以上の手順により、新たな要求R(x)を挿入した要求テーブル131のキューの並び替え処理が終了する。

【0106】次に、要求テーブル131が有する要求Rをハードディスク16に発行する手順を説明する。要求選択部141は、要求テーブル131の先頭から順に、すなわち、順序Pの順番で要求R(r)を取り出す。この時、先頭の要求R(r)のアクセスタイプtype(r)が「同一」または「連続」である場合、R(r)の次の要求R(r<sub>+1</sub>)も続けて取り出す。そして、要求R(r<sub>+1</sub>)についてもアクセスタイプtype(r<sub>+1</sub>)を確認し、アクセスタイプが「独立」となるまで連続して要求Rを取り出す。この連続して取り出すことができる要求Rの数は、アクセスタイプが「同一」のものを除き、上述したデータサイズ(=ds)の上限の場合と同様に、アクセス時間テーブル152が保持しているテーブルに制限される。従って、より大きなデータサイズに対応したアクセス時間テーブル152を保持することで、連続して取り出す要求Rの数よ

並び替え制御部11は、アクセス時間テーブル152が抽出したアクセス時間t{g(x<sub>-n</sub>), g(x)}を要求R(x)のアクセス時間t(x)とし、完了予定時刻e(x)を下記式(1)に基づいて算出する(ステップS308)。

9)。なお、式(2)において完了予定時刻e(x<sub>+n</sub>)'は、要求R(x<sub>+n</sub>)が要求R(x)の挿入前に有していた完了予定時刻である。

り多くすることが可能である。

【0107】要求選択部141は、上記単独の要求R(r)または連続して取り出した要求R(r)~R(r<sub>+1</sub>)に関する一つのコマンドC(s)をコマンド管理テーブル141a(図7を参照)上で生成する。コマンド発行部142は、要求選択部141で生成したコマンドC(s)をハードディスク16に発行する。また、コマンド発行部142は、コマンドC(s)を発行してからハードディスク16における処理完了の応答を受信するまでの実行時間exe(s)を測定する。そして、コマンド発行部142は、実行時間exe(s)を、種類K(s)、データサイズlen(s)、先頭論理アドレスA(s)と共に、アクセス時間管理部153へ通知する。

【0108】また、コマンド発行部142は、応答を受信した結果を要求選択部141に通知する。そして、要求選択部141は、この通知を受けてユーザに対し要求の処理が終了した旨を任意の方法で返送する。この際、要求選択部141は、1つのコマンドC(s)で複数の要求を実行したものについては、各要求ごとの実行結果として返送する。この場合、読み出し要求であれば、読み出したデータを先頭から64kバイト単位に区切り、コマンド管理テーブル141aより各要求のデータ位置dが示すデータを応答として返送する。以上の手順により、要求テーブル131が有する要求Rに対するハードディスク16への処理が終了する。

【0109】最後に、運用中において、アクセス時間テーブル152に格納しているアクセス時間tの更新を行う動作を説明する。なお、このアクセス時間tの更新を行う動作は、上述したアクセス時間tの初期値を設定する際の動作と同様である。アクセス時間算出部153は、上記のようにコマンド発行部142から今実行したコマンドC(s)に関する各情報を入手する。そして、コマンドC(s)の直前に実行したコマンドC(s<sub>-1</sub>)の論理アドレスA(s<sub>-1</sub>)と、論理アドレスA(s)とに基づいて、グループ変換部151によりグループg(s)およびg(s<sub>-1</sub>)を求め、その後、アクセス時

間算出部153は、アクセス時間テーブル152内の種類 $K(s)$ およびデータサイズ $len(s)$ が一致するテーブルについて、アクセス時間 $t\{g(s_{-1}), g(s)\}$ の値を実行時間 $exe(s)$ の値に書き替える。このような処理を、コマンドCを実行することに行い、常に最新のアクセス時間 $t$ を求め、アクセス時間テーブル152内の値を更新する。

【0110】なお、並べ替え制御部11は、新たな要求 $R(x)$ の有無にかかわらず、アクセス時間テーブル152内の値が更新することにより、要求テーブル131の要求の並べ替えを行なうようにしてもよい。

【0111】以上のように、本発明の第1の実施形態に係るディスク制御装置は、要求が入力されるたびに、要求の実行時間を予測し、各要求に指定された遅延時間の限界値以内に要求の実行が完了するように要求を並べ替える。また、要求の実行時間の予測は保持されたアクセス時間を参照するだけであるから、並べ替えの際に予測値を得るために特別な演算を必要としない。従って、アクセス効率がよくなり、かつ、遅延時間の限界値以内に要求を実行することが可能となる。特に、動画や音声データなどのリアルタイム性を必要とするデータにアクセ

$$T = (T_1 - T_2) / (L_1 - L_2) \times L + T_1 / L_1 \times (L_1 - L_2) / (T_1 - T_2) \quad \dots (3)$$

【0114】また、本第1の実施形態では、許容遅延時間 $delay$ は要求 $R$ 内に指定されているとしたが、あらかじめ何通りかの許容遅延時間を用意し、それぞれの許容遅延時間に識別番号を割り付け、要求 $R$ 内に指定する値は、その識別番号でも構わない。この場合、遅延時間管理部12内に、許容遅延時間値とその識別番号の対応表を保持し、完了時刻算出部121は、上記対応表により許容遅延時間値に変換し、完了時刻 $f$ を算出してもよい。また、許容遅延時間をシステム固定の値とし、完了時刻算出部121は、上記固定の値を用いて完了時刻 $f$ を算出してもよい。さらに、要求の種類 $k$ に対応してシステム固定の値を保持し、完了時間算出部121は、要求の種類 $k$ を参照することで許容遅延時間を獲得し、完了時刻 $f$ を算出してもよい。

【0115】また、本第1の実施形態では、連続したブロックにアクセスする要求をひとまとめでして1つのコマンドCで実行することを可能にしていたが、特にこの機能に対応していなくてもよい。すなわち、要求選択部141は、アクセスタイプ $type$ が「連続」の場合の動作を実行しなくてもよい。この場合、上記第1の実施形態では、ハードディスク16にアクセスするデータサイズは、各要求がアクセスするサイズと常に等しくなる。従って、各要求がアクセスするデータ長を固定長のみとし、アクセス時間テーブル152は、固定長に対応した値のみを保持してもよい。同様に、予め各要求がアクセスするデータ長を何通りか決めておき、そのデータ長に応じてアクセス時間テーブル152を用意し、アク

スする場合に、アクセス効率を向上させ、かつ、リアルタイム性を確保することが可能となる。

【0112】なお、本第1の実施形態では、読み出し要求および書き込み要求を扱ったが、これ以外の要求に対しても実行可能である。この場合、追加する要求の種類に対応したアクセス時間テーブルを作成し、追加した要求のアクセス時間を格納すれば良い。また、読み出し要求のみ、または書き込み要求のみでも構わない。

【0113】また、各要求は、全て64kバイトの固定長データをアクセスする要求であったが、要求内にアクセスするデータ長を指定し、要求ごとに可変データ長をアクセスしてもよい。この場合、データ長に対応したアクセス時間テーブルを参照する方法や、2種類のデータ長に対応したアクセス時間テーブルを用意し、該当するデータ長のアクセス時間を先の2種類のアクセス時間を用いて算出する方法がある。具体的には、データ長 $L_1$ のアクセス時間を $T_1$ 、データ長 $L_2$  ( $L_1 < L_2$ ) のアクセス時間を $T_2$ とし、算出するアクセス時間を $T$ 、そのデータ長を $L$ とすると、アクセス時間 $T$ は、例えば一次式を用いて補間し、下記式(3)により求まる。

セス時間 $t$ を求めてもよい。

【0116】また、本第1の実施形態では、ディスク制御装置の運用前にアクセス時間 $t$ の初期化値を求めたが、コマンド生成部143、コマンド発行部142、アクセス時間管理部15と同等の構成を持つ別の装置により取得したアクセス時間 $t$ をアクセス時間テーブル152に設定してもよい。また、その設定値は、ハードディスク16上に記録されており、コマンド生成部143は上記設定値を読み出すコマンドCを生成し、コマンド発行部142を通してハードディスク16より獲得し、アクセス時間テーブル152に格納してもよい。また、コマンド生成部143は、アクセス時間テーブル152からアクセス時間テーブル152内に保持した値を受けとり、その値を書き込みデータとする書き込みコマンドを生成し、コマンド発行部142を経由してハードディスク16上の所定の位置に書き込んでもよい。この書き込んだ値は、コマンド生成部143がこの値を読み出すコマンドCを生成し、コマンド発行部142からコマンドCを発行することで読み出すことが可能である。

【0117】また、本第1の実施形態においては、ハードディスク16以外にアクセス時間 $t$ を書き込みおよび読み出し可能な別の記憶装置を設け、この記憶装置に対するコマンドCを生成するコマンド生成部と、コマンドCを記憶装置に発行し、また応答を受信するコマンド発行部とを新たに設け、アクセス時間 $t$ を上記記憶装置内に書き込むことで保存し、読み出してアクセス時間テーブル152に格納してもよい。

【0118】また、本第1の実施形態において、アクセス時間算出部153は、2つのグループgの順列ごとに、一定数の要求Rの実行時間exeを元にアクセス時間tを更新したが、実行した全ての要求Rについて、一定数ごとに2つのグループgの順列全てについて、アクセス時間tを算出し更新してもよいし、実行した全ての要求Rについて、2つのグループgの順列それぞれについて、一定数ごと要求の実行が完了するたびに要求の実行が完了したグループgの順列のアクセス時間tを算出し更新してもよい。

【0119】さらに、本第1の実施形態において、遅延時間管理部12は、要求ごとに完了時刻fを算出し保持するが、遅延時間管理部12は、要求ごとに許容遅延時間delayを初期値とし、要求の実行が完了するまで許される時間である許容時間を保持し、並べ替えのたびに許容時間を更新し、並べ替え部141は、キューの先頭から要求を挿入する直前の要求までのアクセス時間tの和を算出し、許容時間とアクセス時間aの和を比較し、アクセス時間の和が小さい場合に完了予定時刻eが完了時刻f以下とみなし、並べ替えを行なってもよい。

【0120】(第2の実施形態)図12は、本発明の第2の実施形態に係るディスク制御装置の構成を示すブロック図である。図12において、本第2の実施形態に係るディスク制御装置は、並べ替え制御部21と、遅延時間管理部22と、要求管理部23と、ディスクIO部24と、ハードディスク26とを備える。

【0121】並べ替え制御部21は、ユーザ管理部(図示せず)から処理の要求を受け付ける。そして、並べ替え制御部21は、遅延時間管理部22から与えられる情報に基づいて、シーク距離が短くなるように、要求管理部23が管理する処理待ち要求の並び(キュー)の任意の位置に受け付けた要求を挿入する。なお、ユーザ管理部は、上記第1の実施形態において説明したものと同様であり、以下その説明を省略する。

【0122】遅延時間管理部22は、要求管理部23が管理する要求の許容遅延時間を算出して保持する。要求管理部23は、並べ替え制御部21から連続的に入力される複数の要求をキューとして保持する。なお、要求管理部23が管理するキューの模式図は、図2に示したものと同様である。ディスクIO部24は、要求管理部23のキューから要求を順に読み出し、対応するコマンドをハードディスク26へ出力する。このコマンドは、所定のサイズのデータの読み出しまたは書き込みを行うコマンドである。ハードディスク26は、複数種類のマルチメディアデータを一定のサイズごとに分割して記憶する。

【0123】まず、図13および図14を用いて、本第2の実施形態に係るディスク制御装置を具体的に説明する。図13は、本第2の実施形態に係るディスク制御装置のさらに詳細な構成を示すブロック図である。図14

は、要求テーブル231の構成の一例を示す図である。

【0124】図13を参照して、遅延時間管理部22は、挿入可能要求数算出部221と、平均実行時間算出部222と、平均実行時間保持部223とを備える。また、ディスクIO部24は、要求選択部241と、コマンド発行部242と、コマンド生成部243とを備える。また、要求管理部23は、キューを要求テーブル231の形態で管理している。

【0125】挿入可能要求数算出部221は、並べ替え制御部21から与えられる許容遅延時間delay(x)と、平均実行時間保持部223が保持するコマンドの平均実行時間とから、挿入可能要求数i(x)を算出する。平均実行時間算出部222は、コマンド発行部242から受けた実行時間に基づいて、平均実行時間を算出する。平均実行時間保持部223は、平均実行時間算出部222が算出したコマンドの平均実行時間を保持する。

【0126】要求テーブル231は、上述のように要求を管理するためのテーブルである。要求テーブル231は、図14に示すように、各要求Rごとに種類k、論理アドレスa、挿入可能要求数iおよび禁止フラグflagをそれぞれ記憶する。なお、図14における順序Pは、要求の処理順序、すなわち、要求選択部241が取り出す要求の順序を示している。1がキューの先頭を、最大値nがキューの末尾を表す。このnは、キュー長でもある。

【0127】種類kは、要求が読み出し要求であるか書き込み要求であるかを表す。論理アドレスaは、要求Rによって処理すべきデータのハードディスク26上の先頭位置を示す論理的なアドレスである。挿入可能要求数iは、キュー内において、要求R(x)よりも前に挿入可能な要求Rの数を示す。禁止フラグflagは、フラグONによって要求Rの前に新たな要求を挿入できないこと、または、要求Rの前後でシーク方向が逆になることを表す。

【0128】要求選択部241は、要求テーブル231の先頭から要求を1つ取り出し、ハードディスク26に発行するコマンドCを生成する。コマンド生成部243は、平均実行時間保持部223で保持するコマンドの平均実行時間の初期値を得るためのコマンドCを生成する。コマンド発行部242は、要求選択部241またはコマンド生成部243が生成したコマンドCを受けて、ハードディスク26へ発行する。また、コマンド発行部242は、発行したコマンドCの実行結果をハードディスク26から受信し、要求選択部241に実行結果を通知する。さらに、コマンド発行部242は、コマンド発行から実行結果の受信までの実行時間を測定し、平均実行時間算出部222へ出力する。

【0129】並べ替え部21は、新たに受けとった要求R(x)を、各要求の挿入可能要求数i(x)を越えな



いように、また各要求がアクセスする論理アドレス  $a(x)$  の順序がなるべく一定となるように、並べ替える。

【0130】次に、図15および図16を参照して、本第2の実施形態に係るディスク制御装置が行う処理を具体的に説明する。図15は、平均実行時間保持部223の初期設定手順を示すフローチャートである。図16は、新たな要求  $R(x)$  の挿入処理手順を示すフローチャートである。

【0131】まず、前提として、平均実行時間保持部223に保持する平均実行時間の初期値を求める方法を、図15を参照して説明する。コマンド生成部243は、ハードディスク26上のランダムな論理アドレスから始まる64kバイトのデータにアクセスするコマンド  $C$  を生成する。コマンド  $C$  を生成する際、コマンドの種類  $k$  (要求の種類に相当) についてもランダムに選択する(ステップS401)。

【0132】コマンド発行部242は、コマンド生成部243が生成したコマンド  $C$  を受けとり、ハードディスク26へ発行し、その応答を受信するまでの実行時間を測定する(ステップS402)。平均実行時間算出部222は、コマンド発行部242からコマンド  $C$  の種類  $k$  と実行時間とを取得する。そして、平均実行時間算出部222は、実行時間を種類  $k$  によって分類し、分類ごとに実行時間の和を算出する(ステップS403)。さらに、平均実行時間算出部222は、受けとった実行時間の数を種類  $k$  ごとにカウントする(ステップS403)。そして、平均実行時間算出部222は、ある種類  $k$  において受けとった実行時間の数が予め定めた値(例えば、10000個)に達したかどうかを判断し(ステップS404)、達した場合には、平均実行時間を算出して平均実行時間保持部223に格納する(ステップS405)。以上の処理によって、平均実行時間の初期値が平均実行時間保持部223に保持される。

【0133】次に、並べ替え制御部21が新しい要求  $R(x)$  を受けとった以後の動作を、図16を参照して説明する。並べ替え制御部21は、ユーザ管理部から新たな要求  $R(x)$  を受け取ると、要求  $R(x)$  の許容遅延時間  $delay(x)$  および種類  $k(x)$  を挿入可能要求数算出部221へ送信する。挿入可能要求数算出部221は、平均実行時間保持部223が保持する平均実行時間で許容遅延時間  $delay(x)$  を除算し、挿入可能要求数  $i(x)$  を算出する(ステップS501)。なお、参照する平均実行時間は、要求  $R(x)$  の種類  $k(x)$  と等しいものとする。そして、挿入可能要求数算出部221は、算出した挿入可能要求数  $i(x)$  を並べ替え制御部21へ返送する。

【0134】次に、並べ替え制御部21は、要求テーブル231をキューの最後尾から順にスキャンし、禁止フラグ  $flag$  がONである要求  $R$  が存在するか否かを検

索する(ステップS502)。この検索処理および後述する比較処理の対象は、要求テーブル231に存在する要求のみならず、直前に要求選択部241によって取り出された要求をも含まれる。ステップS502において該当する要求  $R(j)$  が見つかった場合、並べ替え制御部21は、その要求  $R(j)$  から最後尾までのすべての要求  $R$  の論理アドレス  $a$  と、要求  $R(x)$  の論理アドレス  $a(x)$  との位置を比較する(ステップS503、S505)。ステップS502において該当する要求  $R$  が見つからなかった場合、並べ替え制御部21は、すべての要求  $R$  の論理アドレス  $a$  と、要求  $R(x)$  の論理アドレス  $a(x)$  との位置を比較する(ステップS504、S505)。

【0135】上記ステップS505において論理アドレスの位置を比較した結果、論理アドレス  $a(x)$  が比較した論理アドレスの間に位置する場合は、論理アドレス  $a$  の順序を乱さない(昇順または降順となる)位置に要求  $R(x)$  を挿入する(ステップS506)。一方、ステップS505において論理アドレスの位置を比較した結果、論理アドレス  $a(x)$  が比較した論理アドレスの間に位置しない場合は、キューの最後尾に要求  $R(x)$  を挿入する(ステップS507)。この時、論理アドレス  $a(x)$  が最後尾の要求  $R$  の論理アドレス  $a$ 、または挿入位置直前の要求  $R$  の論理アドレス  $a$  の順序と異なる場合には、最後尾または挿入位置直前の要求  $R$  の禁止フラグ  $flag$  をONに設定する。

【0136】次に、並び替え制御部21は、要求  $R(x)$  を挿入した位置より前に並んでいる要求  $R$  の数を計数し、挿入可能要求数  $i(x)$  から当該計数した要求  $R$  の数を減算する(ステップS508)。そして、並び替え制御部21は、要求  $R(x)$  の次の要求から最後尾の要求までの挿入可能要求数  $i(x_{n+1})$  からそれぞれ1を減算する(ステップS509)。その結果、挿入可能要求数  $i(x_{n+1})$  の値が0となった要求  $R(x_{n+1})$  については、禁止フラグ  $flag(x_{n+1})$  をONに設定する(ステップS510)。以上で要求  $R(x)$  の並び替えが終了する。

【0137】次に、要求テーブル231が有する要求  $R$  をハードディスク26に発行する手順を説明する。要求選択部241は、要求テーブル231の先頭から順に、すなわち、順序  $P$  の順番で要求  $R$  を取り出す。要求選択部241は、取り出した要求  $R$  に関するコマンド  $C$  を生成する。コマンド  $C$  には、種類  $k$ 、論理アドレス  $a$ 、データサイズが含まれる。コマンド発行部242は、要求選択部241で生成したコマンド  $C$  をハードディスク26に発行する。また、コマンド発行部242は、コマンド  $C$  を発行してからハードディスク26における処理完了の応答を受信するまでの実行時間  $exe$  を測定する。そして、コマンド発行部242は、実行時間  $exe$  を種類  $K$  と共に、実行時間算出部222へ通知する。また、

コマンド発行部242は、応答を受信した結果を要求選択部241に通知する。そして、要求選択部241は、この通知を受けてユーザに対し要求の処理が終了した旨を返送する。以上の手順により、要求テーブル231が有する要求Rに対するハードディスク16への処理が終了する。なお、言うまでもなく、1つの要求の処理が終わるごとに要求テーブル231に存在する各要求の挿入可能要求数*i*の値は、1つずつ減算されていく。

【0138】最後に、運用中において、平均実行時間保持部223に保持している平均実行時間の更新を行う動作を説明する。なお、この平均実行時間の更新を行う動作は、上述した平均実行時間の初期値を設定する際の動作と同様である。平均実行時間算出部222は、上記のようにコマンド発行部242から実行したコマンドCの実行時間および種類*k*を入手する。そして、平均実行時間算出部222は、実行時間を受けとった種類によって分類し、分類ごとの実行時間の和を算出する。また、実行時間の数を種類*k*ごとにカウントする。そして、平均実行時間算出部222は、ある種類*k*において、受けとった実行時間*e<sub>x</sub>e*の数が予め定めた値に達した場合、平均実行時間を算出し、平均実行時間保持部223へ格納する。このような処理を、コマンドCを実行することに行い、常に最新のアクセス時間*t*を求め、平均実行時間保持部223内の値を更新する。

【0139】以上のように、本発明の第2の実施形態に係るディスク制御装置は、要求が入力されるたびに、要求が保持する遅延時間の限界値から挿入可能要求数を求め、挿入可能要求数よりも多くの要求をその要求の前に挿入しないように、要求を並べ替える。従って、アクセスの効率が良くなり、かつ、遅延時間の限界値以内に要求の実行を完了することができる。

【0140】なお、本第2の実施形態では、読み出し要求と書き込み要求を扱ったが、これ以外に扱う要求を追加してもよい。この場合、平均実行時間保持部223は、追加した要求の種類*k*に対応する平均実行時間を保持し、挿入可能要求数*i*を平均実行時間から算出する。

【0141】また、本第2の実施形態では、運用中のコマンドの実行時間を測定することで平均実行時間を更新したが、平均実行時間の更新を行わなくてもよい。また、平均実行時間の初期値を測定したが、この初期値は、本第2の実施形態と同様の構成を持つ他の装置から取得し、平均実行時間保持部223に保持するようにしてもよい。

【0142】また、本第2の実施形態において、コマンド生成部243は、平均実行時間をハードディスク26上の所定の位置に書き込むコマンドと、ハードディスク26から読み出すコマンドとを生成し、書き込みコマンドによりハードディスク26上に平均実行時間を保存し、また読み出しコマンドによりハードディスク26から平均実行時間を取得し、平均実行時間保持部223に

格納してもよい。

【0143】また、本第2の実施形態においては、平均実行時間を記憶可能なハードディスク26以外の記憶装置と、その記憶装置に対する書き込みおよび読み出しコマンドを生成するコマンド生成部と、生成したコマンドを発行するコマンド発行部とを備え、書き込みコマンドにより記憶装置に平均実行時間を保存し、読み出しコマンドにより記憶装置より平均実行時間を読み出し、平均実行時間保持部223に格納するようにしてもよい。

【0144】さらに、本第2の実施形態において、要求選択部241は、常に1つの要求のみを選択したが、ハードディスク26上で同じまたは連続したデータ領域にある2つの要求を1つにまとめ、ハードディスク26にコマンドを発行してもよい。この場合、要求選択部241は、要求テーブル231から要求を取り出す際に、取り出した要求および要求テーブル231の先頭の要求について、論理アドレス*a*、要求の種類*k*、挿入可能要求数*i*を参照する。要求の種類*k*が共に等しく、論理アドレス*a*が等しいか64kバイトだけ異なる値を示し、論理アドレス*a*がより小さい値をもつ要求の挿入可能要求数*i*が0以外であれば、さらにその要求を取り出す。この操作を繰り返し、ハードディスク26に取り出した要求をまとめ、1つのコマンドとして発行してもよい。

【0145】ところで、上述したように本発明の第1および第2の実施形態に係るディスク制御装置は、1つのコマンドに対するディスク装置（ハードディスク16、26）での実行時間を測定し、この時間に基づいて最適な要求の並べ替えを行う。しかし、キャッシュ機能を有しているディスク装置において、キャッシュ機能を働かせている場合、任意のコマンドに対して、ディスク装置に対して行われる読み出し／書き込み動作ではなく、キャッシュに対して行われる読み出し／書き込み動作の応答が返ってくる可能性がある。このため、キャッシュ機能を働かせている場合には、ディスク装置での実行時間を正確に測定することができなくなるといことが考えられる。この対応としては、単にディスク装置のキャッシュ機能を使用しなければよいのである。キャッシュ機能を使用しなくする方法としては、次の2つの方法がある。1つはキャッシュ機能の設定自体をOFFにする方法である。もう1つは、読み出し／書き込み要求内のFUA (Force Unit Access) ビットを指定して要求を発行する方法である。このFUAビットとは、強制的な装置へのアクセスを指定するためのビットであり、指定するとキャッシュ機能の設定をOFFにした場合と同じ動作となる。このように、キャッシュ機能をOFF設定にする、またはFUAビットを指定して要求を発行することにより、いかなるディスク装置に対しても上述した本発明の有用な効果を奏することが可能となる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態に係るディスク制御装

置の構成を示すブロック図である。

【図2】要求管理部13が管理するキューの模式図を示す図である。

【図3】本発明の第1の実施形態に係るディスク制御装置のさらに詳細な構成を示すブロック図である。

【図4】ハードディスク16におけるブロック分割およびグループ分けの一例を示す図である。

【図5】完了時刻テーブル122の構成の一例を示す図である。

【図6】要求テーブル131の構成の一例を示す図である。

【図7】コマンド管理テーブル141aの構成の一例を示す図である。

【図8】アクセス時間テーブル152の構成の一例を示す図である。

【図9】アクセス時間テーブル152の初期設定手順を示すフローチャートである。

【図10】並び替え制御部11における新たな要求R(x)の挿入処理手順を示すフローチャートである。

【図11】図10のステップS2.14の算出処理手順を示すフローチャートである。

【図12】本発明の第2の実施形態に係るディスク制御装置の構成を示すブロック図である。

【図13】本発明の第2の実施形態に係るディスク制御装置のさらに詳細な構成を示すブロック図である。

【図14】要求テーブル231の構成の一例を示す図である。

【図15】平均実行時間保持部223の初期設定手順を示すフローチャートである。

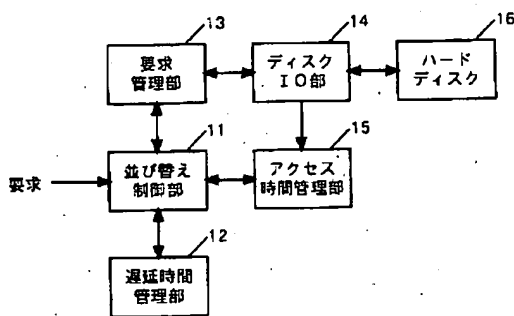
【図16】並び替え制御部21における新たな要求R(x)の挿入処理手順を示すフローチャートである。

【図17】従来のディスク制御装置の構成を示すブロック図である。

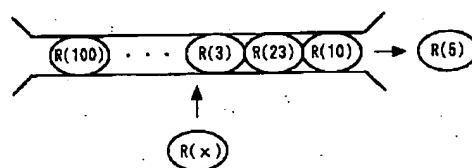
【符号の説明】

11, 21, 301…並び替え制御部  
12, 22…遅延時間管理部  
13, 23, 303…要求管理部  
14, 24…ディスクIO部  
15…アクセス時間管理部  
16, 26, 306…ハードディスク  
121…完了時刻算出部  
122…完了時刻テーブル  
131, 231…要求テーブル  
141, 241…要求選択部  
141a…コマンド管理テーブル  
142, 242…コマンド発行部  
143, 243…コマンド生成部  
151…グループ変換部  
152…アクセス時間テーブル  
153…アクセス時間算出部  
221…挿入可能要求数算出部  
222…平均実行時間算出部  
223…平均実行時間保持部  
302…許容遅延データ管理部

【図1】



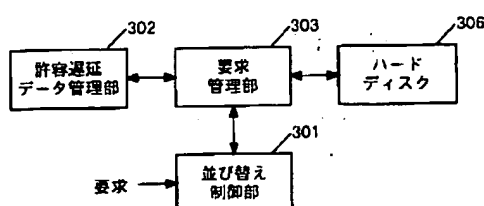
【図2】



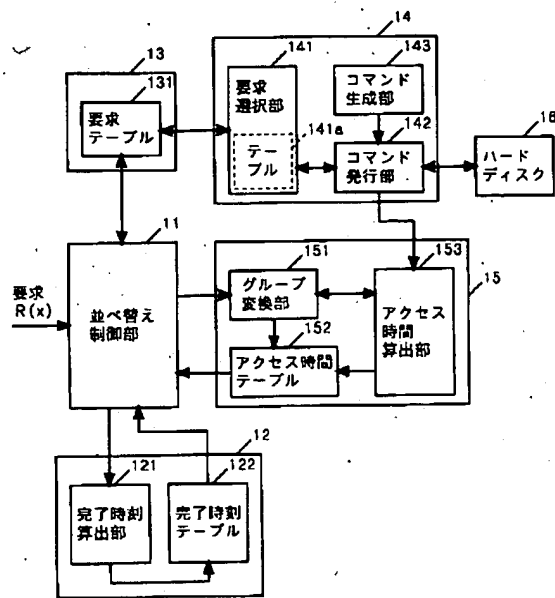
【図5】

要求 R	完了時刻 f
R(1)	f(1)
R(2)	f(2)
R(3)	f(3)
⋮	⋮
R(n)	f(n)

【図17】



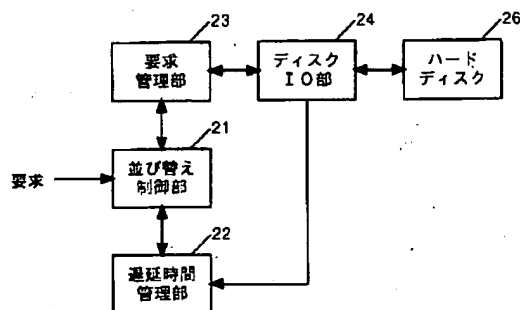
【図3】



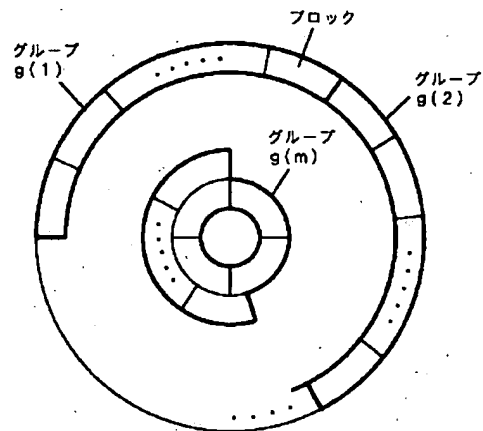
【図6】

順序 P	要求 R	種類 k	論理 アドレス a	アクセス 時間 t	完了予定 時刻 e	禁止 フラグ flag	アクセス タイプ type
1	R(1)	k(1)	a(1)	t(1)	e(1)	flag(1)	type(1)
2	R(2)	k(2)	a(2)	t(2)	e(2)	flag(2)	type(2)
3	R(3)	k(3)	a(3)	t(3)	e(3)	flag(3)	type(3)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	R(n)	k(n)	a(n)	t(n)	e(n)	flag(n)	type(n)

【図12】



【図4】



【図7】

コマンド C	種類 k	先頭論理 アドレス A	サイズ len	要求 R	位置 d
C(1)	K(1)	A(1)	len(1)	R(1)	d(1)
				R(2)	d(2)
				R(3)	d(3)
C(2)	K(2)	A(2)	len(2)	R(4)	d(4)
				R(5)	d(5)
				⋮	⋮

【図8】

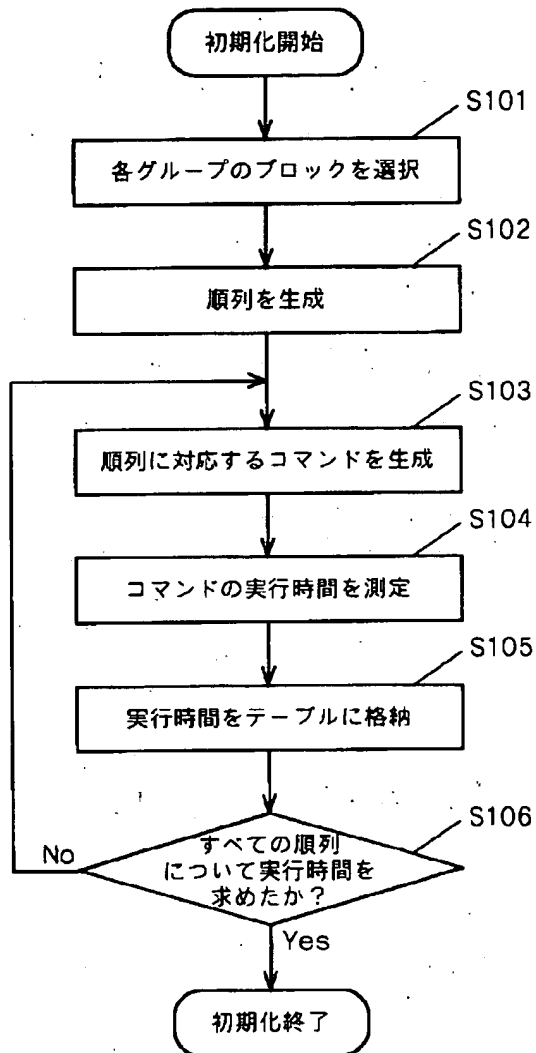
READ 64kB	要求R(x)がアクセスするグループ					
	1	2	3	...	m	
グループ 1	t{1,1}	t{1,2}	t{1,3}	...	t{1,m}	
	t{2,1}	t{2,2}	t{2,3}	...	t{2,m}	
	t{3,1}	t{3,2}	t{3,3}	...	t{3,m}	
	⋮	⋮	⋮	⋮	⋮	
	t{m,1}	t{m,2}	t{m,3}	...	t{m,m}	

(a)

READ 64kB	要求R(x)がアクセスするグループ					
	1	2	3	...	m	
グループ 1	t{1,1}	t{1,2}	t{1,3}	...	t{1,m}	
		t{2,2}	t{2,3}	...	t{2,m}	
			t{3,3}	...	t{3,m}	
			⋮	⋮	⋮	
					t{n,m}	

(b)

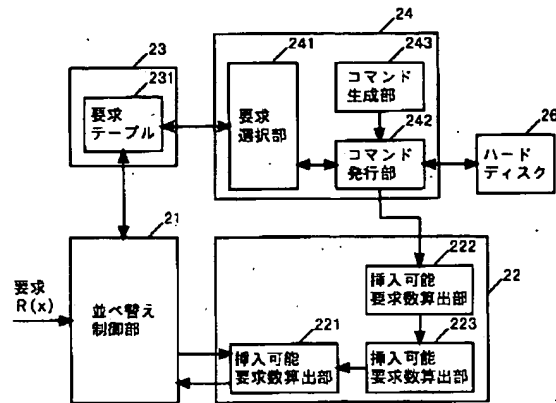
【図9】



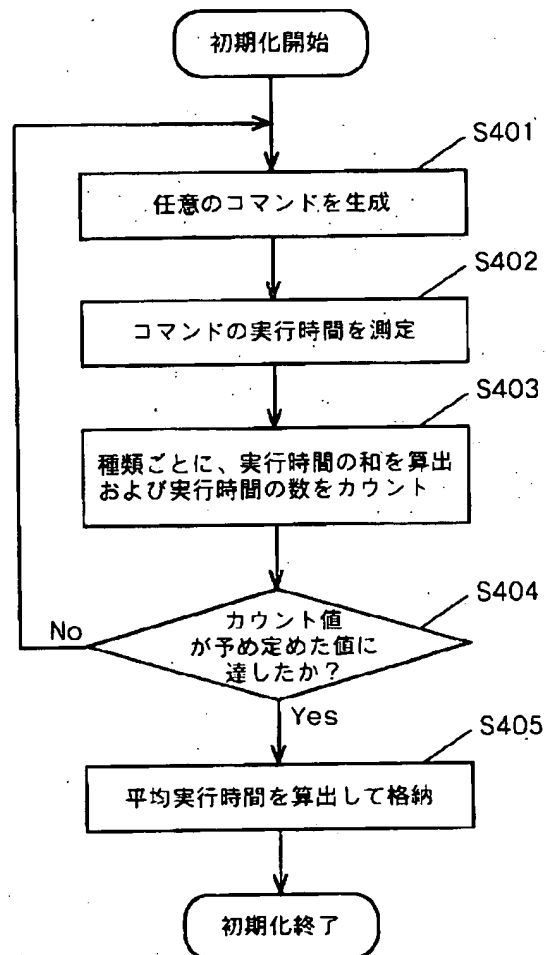
【図14】

順序 P	要求 R	種類 k	論理 アドレス a	挿入可能 要求数 i	禁止 フラグ flag
1	R(1)	k(1)	a(1)	i(1)	flag(1)
2	R(2)	k(2)	a(2)	i(2)	flag(2)
3	R(3)	k(3)	a(3)	i(3)	flag(3)
⋮	⋮	⋮	⋮	⋮	⋮
n	R(n)	k(n)	a(n)	i(n)	flag(n)

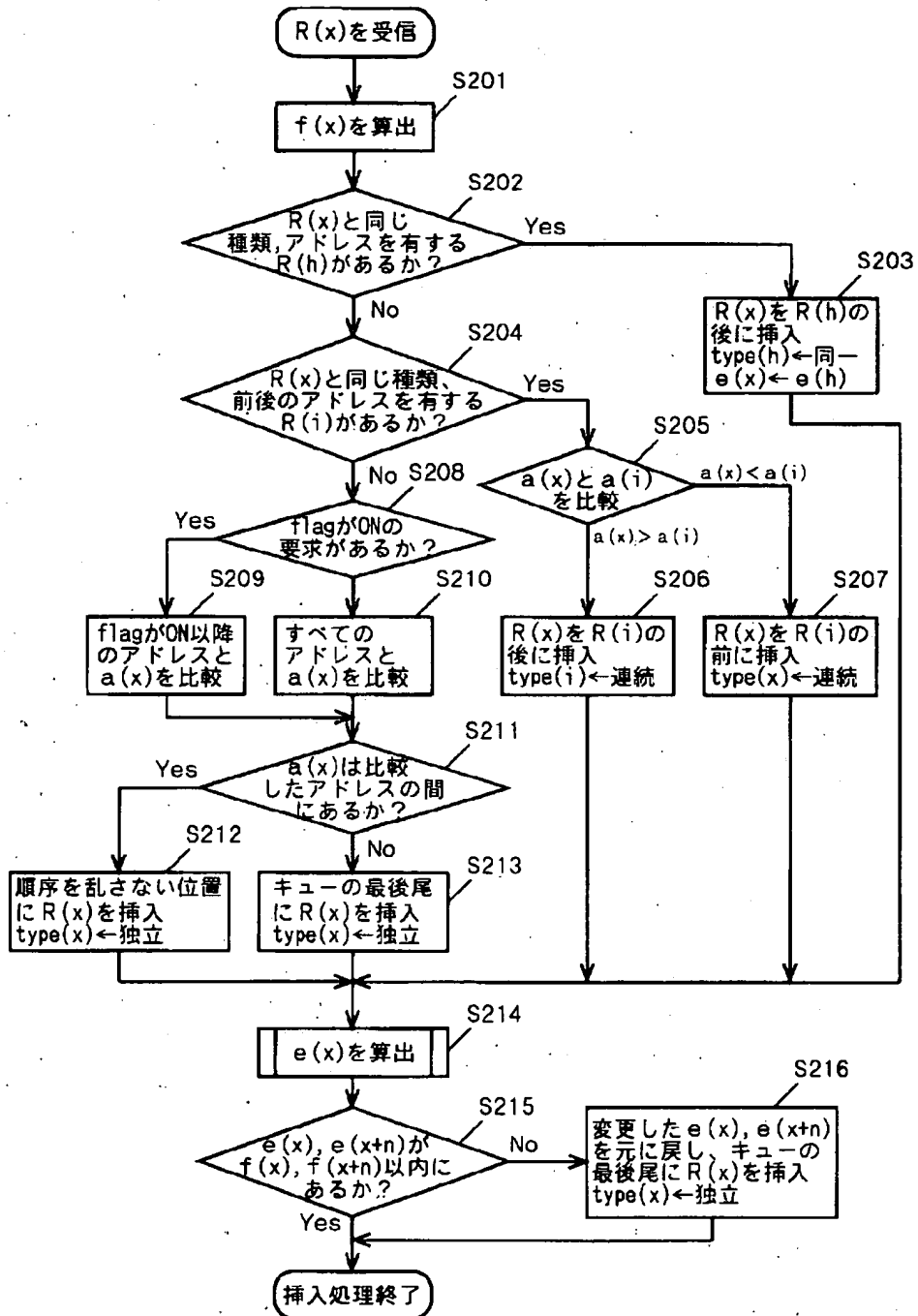
【図13】



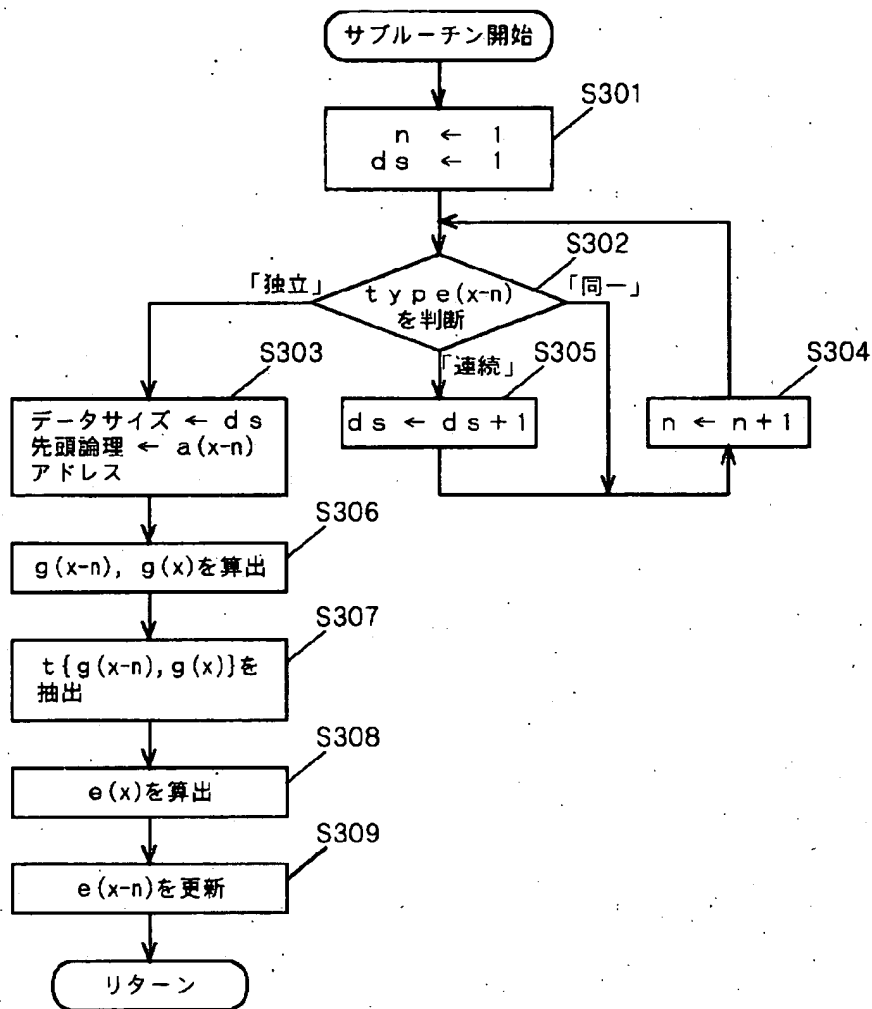
【図15】



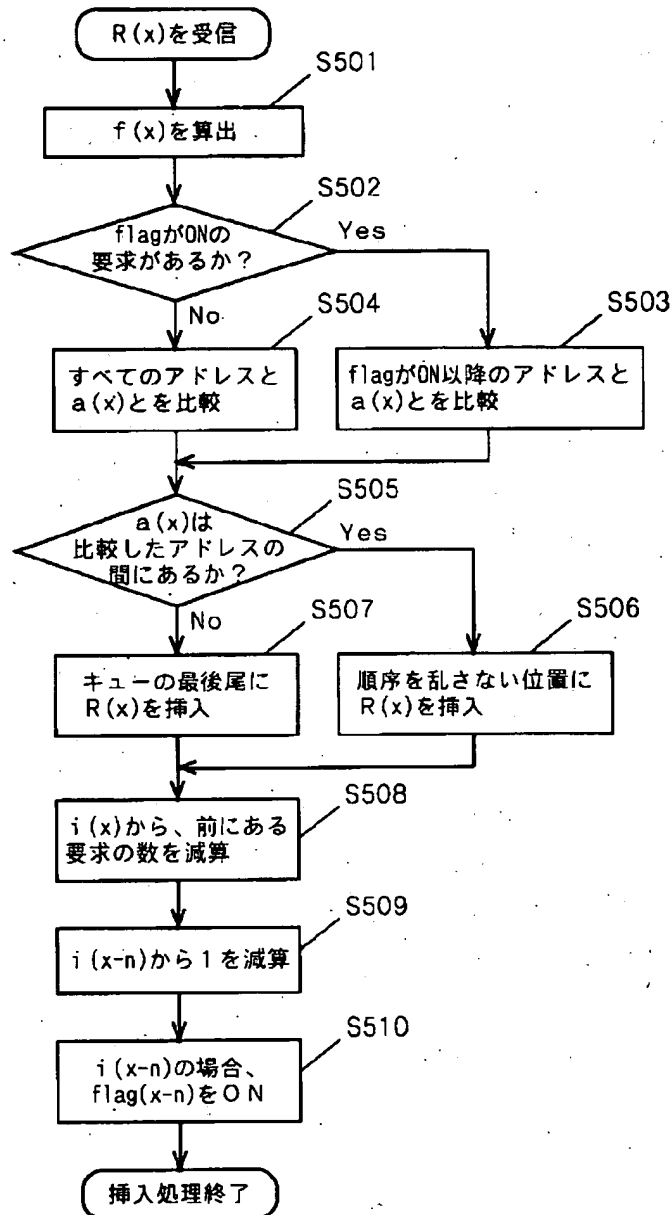
【図10】



【図11】



【図16】



フロントページの続き

(72)発明者 伊藤 由起子  
大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

Fターム(参考) 5B065 BA01 CA12 CC06 CC08 CH15  
EK10